

MODELING TONAL TENSION

FRED LERDAHL
Columbia University

CAROL L. KRUMHANSL
Cornell University

THIS STUDY PRESENTS AND TESTS a theory of tonal tension (Lerdahl, 2001). The model has four components: prolongational structure, a pitch-space model, a surface-tension model, and an attraction model. These components combine to predict the rise and fall in tension in the course of listening to a tonal passage or piece. We first apply the theory to predict tension patterns in Classical diatonic music and then extend the theory to chromatic tonal music. In the experimental tasks, listeners record their experience of tension for the excerpts. Comparisons between predictions and data point to alternative analyses within the constraints of the theory. We conclude with a discussion of the underlying perceptual and cognitive principles engaged by the theory's components.

Received March 16, 2006, accepted October 4, 2006.

Key words: tonal tension, attraction, pitch space, prolongational structure, multiple regression

THE EBB AND FLOW OF tension is basic to the musical experience and has long been of interest in music theory and criticism (Berry, 1976; Hindemith, 1937; Kurth, 1920; Rothfarb, 2002; Schenker, 1935; Zuckerkandl, 1956). It appears to have a direct link to musical affect (Krumhansl, 1996, 1997), and it shapes not only the listening experience but also aspects of musical performance (Palmer, 1996).

Building on the prolongational component in Lerdahl and Jackendoff (1983; hereafter *GTTM*), Lerdahl (2001; hereafter *TPS*) developed a formal model of tonal tension and the related concept of tonal attraction. The model generates quantitative predictions of tension and attraction for the sequence of events in any passage of tonal music. Earlier empirical studies have shown promising connections between the model's predictions and

participants' responses (Bigand, Parncutt, & Lerdahl, 1996; Cuddy & Smith, 2000; Krumhansl, 1996; Lerdahl & Krumhansl, 2004; Palmer, 1996; Smith & Cuddy, 2003). Our purpose here is to provide a comparatively comprehensive empirical treatment and analysis of the model's predictions over a range of musical styles.

By "tonal tension" we mean not an inclusive definition of musical tension, which can be induced by many factors, such as rhythm, tempo, dynamics, gesture, and textural density, but the specific sense created by melodic and harmonic motion: a tonic is relaxed and motion to a distant pitch or chord is tense; the reversal of these motions causes relative relaxation. Because tonal tension is a uniquely musical phenomenon (unlike such factors as fluctuations in loudness, speed, or contour), it is perhaps the most crucial respect in which music tenses and relaxes. This study sets aside other kinds of musical tension and focuses on tonal tension.

The sense of tonal tension and relaxation can also be expressed as "stability and instability" or even "consonance and dissonance." These pairs of terms have somewhat different shades of meaning. "Dissonance" refers first of all to a sensory property that is studied in the psychoacoustic literature. In a traditional music-theoretic context, it refers to intervallic combinations that require particular syntactic treatment, such as the passing tone and the suspension. Intervals that are musically dissonant usually correspond to intervals that are psychoacoustically dissonant. "Instability" has cognitive or conceptual meaning beyond psychoacoustic effects. Theorists such as Riemann (1893), Schenker (1935), and Schoenberg (1911) extend musical dissonance from a surface characteristic to abstract levels. One may speak of a composed-out passing tone that is harmonized at the surface, or of a subsidiary tonal region that is conceptually dissonant in relation to the tonic (Rosen, 1972). Schoenberg (1975) asserts that the goal of a tonal composition, after its initial destabilization, is to reestablish stability.

The term "tension," as employed here, refers both to sensory dissonance and to cognitive dissonance or instability; similarly, "relaxation" refers to sensory consonance and to cognitive consonance or stability. The expression

“tension and relaxation” also has the advantage of invoking physical motion and exertion beyond a specifically musical function. Everyone experiences physical tension and relaxation, and it is common to extend the terms to mental and emotional terrains as well. Consequently, it is relatively straightforward to ask experimental participants to respond to degrees of tension and relaxation and thereby elicit consistent interpersonal responses (see Krumhansl, 1996).

The *TPS* model also develops an attraction component. The term “attraction” refers to the intuition that melodic or voice-leading pitches tend toward other pitches in greater or lesser degrees. Bharucha (1984) refers to melodic anchoring; Larson (2004; Larson & VanHandel, 2005) speaks of musical forces; Margulis (2005), Meyer (1956), and Narmour (1990) couch attraction in terms of melodic expectation or implication. Attraction can also be seen as a kind of tension: the more attracted a pitch is to another pitch, such as the leading tone to the tonic, the more the listener experiences the tension of anticipation. This kind of tension contrasts with the tension of instability. The leading tone is less stable than the tonic, but its expectancy-tension (to use Margulis’s expression) is much greater than that of the tonic. That is, the leading tone strongly “wants” to resolve to the tonic; but the tonic pitch, being the point of maximal stability, expresses comparatively little urge to move to the leading tone or to any other pitch.

To summarize, our concern is with three kinds of tonal tension: the sensory dissonance of certain intervallic combinations, harmonic and regional stability/instability in relation to a governing tonic, and melodic attraction as a projection of expectancy-tension.

Overview of the Tension Model

The four components listed in Figure 1 are required for a quantitative theory of tonal tension. First, there must be a representation of the hierarchical event structure in a musical passage. Adapting a traditional music-theoretic term, we call this component prolongational structure. Second, there must be a model of tonal pitch space and all distances within it. Tonal pitch space is the cognitive schema whereby listeners have tacit long-term knowledge, beyond the patterns within any particular piece, of the distances of pitches, chords, and tonal regions from one another. Third, there must be a treatment of surface or sensory dissonance. This measure is largely psychoacoustic: the interval of a seventh is more dissonant than a sixth, and so on. Fourth, there must be a model of melodic or voice-leading attractions. Listeners

1. A representation of hierarchical (prolongational) event structure.
2. A model of tonal pitch space and all distances within it.
3. A treatment of surface (largely psychoacoustic) dissonance.
4. A model of voice-leading (melodic) attractions.

FIGURE 1. Four components necessary for a quantitative theory of tonal tension.

experience the relative pull of pitches toward other pitches in a tonal context.

Let us review these four components, starting with prolongational structure. (This exposition summarizes material in *TPS*.) *GTTM* addresses prolongational organization not as an aesthetic ideal, as in Schenkerian analysis, but as a psychological phenomenon describable by nested patterns of tension and relaxation. Tension depends on hierarchical position: a tonic chord in root position is relaxed; another chord or region is relatively tense in relation to the tonic; a nonharmonic tone is tense in relation to its harmonic context. This component assigns prolongational structure by a cognitively motivated rule system that proceeds from grouping and meter through time-span segmentation and reduction. These steps are necessary because prolongational connections depend not only on degrees of pitch similarity and stability but also on the rhythmic position of events.

To represent an event hierarchy, the prolongational component employs a tree notation. Here it will suffice to refer to branchings stripped of the node types employed in *GTTM* and *TPS*. Right branches stand for a tensing motion (or departure), left branches for a relaxing motion (or return). The degree of tension or relaxation between two events depends on the degree of continuity between them. If two events that connect are the same or similar, there is little change in tension. If they are different, there is more change in tension. Figure 2 shows an abstract tension pattern: at a local level, Event 1 tenses into Event 2, Event 3 relaxes into Event 4, and

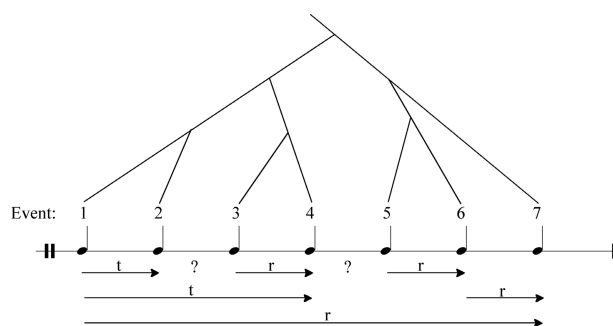


FIGURE 2. Tension (t) and relaxation (r) represented by a tree structure.

Event 5 relaxes into Event 6; at larger levels, Event 1 tenses into Event 4, Event 6 relaxes into Event 7, and Event 1 relaxes into Event 7. Notice that this representation says nothing about the tension relationship between Events 2 and 3 or Events 4 and 5. More seriously, it does not quantify the amount of tension or relaxation. It merely says that if two events are connected, one is relatively tense or relaxed in relation to the other.

Further progress in the evaluation of tension depends on the second component listed in Figure 1, a model of tonal pitch space. A well-known finding in music psychology is that listeners' judgments about the distances of pitches, chords, and regions (or keys) from a given tonic form consistent patterns (Bharucha & Krumhansl, 1983; Krumhansl, 1990, hereafter *CFMP*; Krumhansl & Kessler, 1982). These results have been replicated in several ways, using different input materials, participants with varied training, and different task instructions. When submitted to multidimensional scaling, the empirical data are represented as geometrical structures in which spatial distance corresponds to cognitive distance. The regular geometry found for regions (Krumhansl & Kessler, 1982) corresponds to musical spaces proposed earlier by music theorists (Schoenberg, 1954; Weber, 1817-21).

It is striking that listeners share a complex mental schema of the mutual distances of pitches, chords, and regions. But how is this empirical result to be accounted for? Several researchers have proposed explanatory frameworks: *CFMP* through sensitivity to statistical frequency of tone onsets or durations; Bharucha (1987) through neural-net modeling; Parncutt (1989) through psychoacoustic modeling. A fourth approach, which is complementary to the others, has been to develop a music-theoretic formal model of tonal pitch space that correlates with the empirical data and that unifies the treatment of pitches, chords, and keys within a single framework (Lerdahl, 1988; *TPS*). The model begins with the basic space in Figure 3, set to I/C. Regions are designated in boldface type, with upper-case letters for major keys and lower-case letters for minor keys. The numbers in familiar pitch-class notation signify either pitches or pitch classes, depending on context. The basic space is

(a) octave (root) level:	0											(0)	
(b) fifths level:	0				7							(0)	
(c) triadic level:	0		4		7							(0)	
(d) diatonic level:	0	2	4	5	7	9	11					(0)	
(e) chromatic level:	0	1	2	3	4	5	6	7	8	9	10	11	(0)

FIGURE 3. Diatonic basic space, set to I/C (C = 0, C# = 1, . . . B = 11).

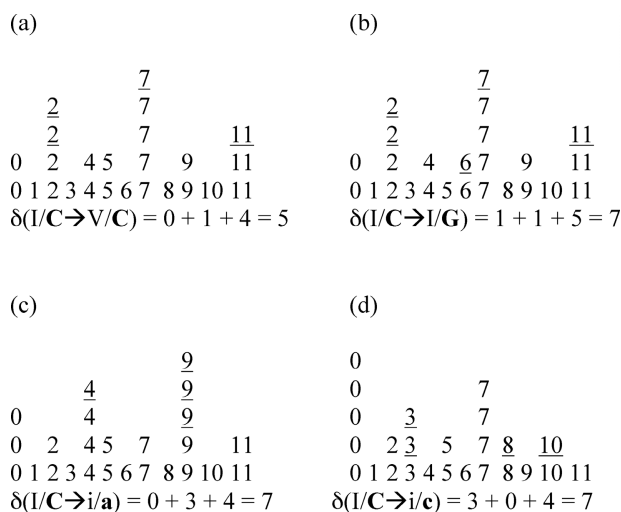
hierarchical in that if a pitch class is stable at one level, it repeats at the immediately superordinate level. The diatonic scale is built from members of the chromatic scale and the triad from members of the diatonic scale. The triad itself has an internal hierarchy, with the fifth more stable than the third and the root as the most stable element. The shape of this structure corresponds to the major-key tone profile in *CFMP* and can be viewed as an idealized form of it.

Transformations of the basic space measure the distance from any chord in any region to any other chord within the region or to any chord in any other region. The space shifts by means of a diatonic chord distance rule in which the distance from chord *x* to chord *y* equals the sum of three variables, as shown in the abbreviated statement of the rule in Figure 4. Computational details aside, the factors involved are the degree of recurrence of common tones and the number of moves along two cycles of fifths, one for triads over the diatonic collection and the other for the diatonic collection over the chromatic collection.

Figure 5 illustrates some basic-space configurations with their distance calculations from I/C (see Figure 3). The underlined numbers signify new pitch classes in the new configuration (variable *k* in the rule). The distance from I/C to V/C in Figure 5a is accomplished by staying in the same diatonic collection (*i*), moving the chord once up the diatonic cycle of fifths (*j*), and counting the resultant noncommon tones (*k*). The distance from I/C to I/G in Figure 5b is two units greater, even though the two chords are the same. In the latter, there is also a cycle-of-fifths move at the scale level (*i*), causing an extra noncommon tone at that level (*k*). Motion between major and minor chords arises not by a transformation but is a by-product of moves along a cycle of fifths.

Diatonic chord distance rule: $\delta(x \rightarrow y) = i + j + k$, where $\delta(x \rightarrow y)$ = the distance between chord *x* and chord *y*; *i* = the number of moves on the cycle of fifths at level (d); *j* = the number of moves on the cycle of fifths at levels (a-c); *k* = the number of non-common pitch classes in the basic space of *y* compared to those in the basic space of *x*.

FIGURE 4. The rule for calculating the distance between triads in diatonic space.

FIGURE 5. Illustrations of δ .

Thus the distance from I/C to i/a in Figure 5c is reached by staying in the same diatonic collection (*i*) and moving the chord three times up the diatonic cycle of fifths (*j*), producing four noncommon tones (*k*). The distance from I/C to i/c in Figure 5d, in contrast, involves moving the scale three times down the chromatic cycle of fifths (*i*). With no change of chord root (*j*), the third of the chord becomes minor. Again there are four noncommon tones (*k*).

When mapped geometrically, the distances (δ) from triad to triad within a key exhibit a regular pattern, with the diatonic cycle of fifths arrayed on the vertical axis and the diatonic cycle of minor thirds on the horizontal axis. Figure 6a displays this pattern along with distances from the tonic triad to the other triads within the key. Regional space—that is, distances from a given tonic triad to other tonic triads—shows a similar pattern, with the chromatic cycle of fifths on the vertical axis and the minor-third cycle on the horizontal axis. Figure 6b gives a portion of regional space along with the distance values. If these chordal and regional patterns are extended, both Figures 6a and 6b form toroidal structures. (Figure 6b corresponds to a multidimensional solution developed from empirical data in Krumhansl and Kessler, 1982; also see *CFMP*.)

Pitch-space distances are input to prolongational structure via the principle of the shortest path. The idea is that listeners construe their understanding of melodies and chords in the most efficient way; in other words, they interpret events in as stable and compact a space as possible. For example, if one hears only the melodic progression $C \rightarrow E$, the most stable interpretation is as $\hat{1} \rightarrow \hat{3}$ in C , for C and E are then in an optimally stable location in

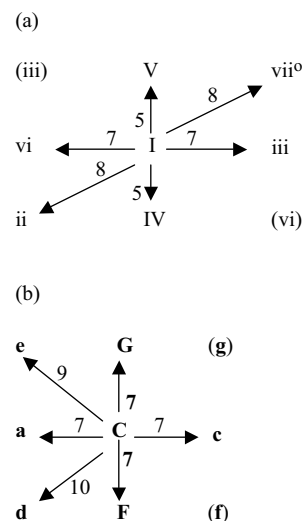


FIGURE 6. Portions of (a) chordal space within a region; and (b) regional space, along with values calculated by δ .

diatonic basic space. A slightly less preferred alternative is as $\hat{3} \rightarrow \hat{5}$ in **a**; still less preferred would be $\hat{4} \rightarrow \hat{6}$ in **G**; and so forth. Similarly, a G major chord heard in a C context is likely to be heard by the shortest path as V/C rather than, say, by longer paths to I/G or iii/e.

Figure 7 illustrates the use of the principle of the shortest path in a derivation of the prolongational structure for the final phrase of the Bach chorale, "Christus, der ist mein Leben." (Later on we discuss the entire chorale; also see the extensive analysis in *TPS*, chapter 1.) Let us assume that the phrasal boundaries and metrical grid have already been assigned. As a first step, automatic segmentation rules carve the music into nested rhythmic units so that each event is assigned to a time-span segment. Second, at the quarter-note time-span level at the bottom of the graph, nonharmonic tones are reduced out, the cadence (marked *c*) is designated, and tonic orientation is established by shortest-path measurement. The opening F major triad is taken to be the tonic because the distance to itself is 0 ($\delta[F \rightarrow F] = 0$), whereas distances to other possible tonics would be greater. All the subsequent events take place within F. Third, events at the quarter-note level come up for comparison at the half-note level. In each case, the most stable event is selected for comparison in the next larger span, where stability is defined in terms of the distance to another available event. Thus the opening I is compared within span *a* to vii^{o6} and I^6 , and ii_2^6 is compared within span *b* to the V. In span *a*, the opening I is selected over vii^{o6} because, unlike vii^{o6} , the distance of I to the tonic is 0; I wins over I^6 because its root is in the bass. In span *b*, the V is chosen because it

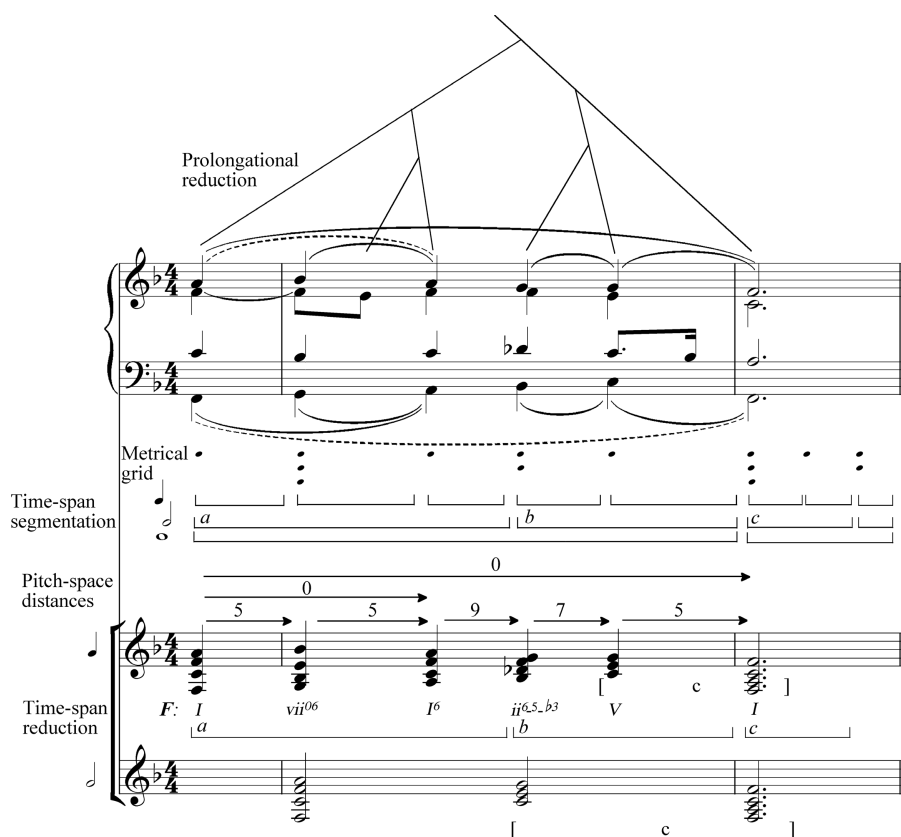


FIGURE 7. Derivation of the prolongational structure of the final phrase of the Bach chorale, "Christus, der ist mein Leben."

is part of the cadence. In span *c*, the only choice is the final *I*. Thus the half-note time-span level yields *I*-*V*-*I*.

The time-span hierarchy then forms the input to the prolongational tree, moving from global to local levels. The distances between available global events are $\delta(\text{opening } I \rightarrow \text{final } I) = 0$ and $\delta(V \rightarrow I) = 5$. The first option wins because its path is shorter: the opening *I* branches directly to the closing *I*, and within that context *V* branches to the final *I*. At a more local level, in the first part of the phrase $\delta(I \rightarrow I^6) = 0$ and $\delta(I \rightarrow vii^{06}) = 5$ (counting, as is customary, vii^{06} as an incomplete dominant), so I^6 attaches to *I*; within the context of *I* – I^6 , vii^{06} branches to I^6 . Finally, ii_5^6 lies between I^6 and *V*. $\delta(ii_5^6 \rightarrow I^6) = 9$ and $\delta(ii_5^6 \rightarrow V) = 7$, so ii_5^6 attaches to the more proximate *V*. As a visual aid, the slurs between events in the music duplicate the relationships described in the tree.

Supplementary to the principle of the shortest path is a second factor in the derivation of prolongational reductions, the principle of good form, which encourages optimal patterns of tension and relaxation. This second principle breaks down into three conditions. First is the recursion constraint, in which successive

right or left branches are preferable to unconnected right or left branches. Thus there is pressure to assign the first instance of Figure 8a rather than the second. Second is the balance constraint, in which the number of right and left branches approaches equality. Thus the first instance in Figure 8b is preferred over the second. Third is normative structure, in which there is a preference for at least one right branch leading off the structural beginning of the phrase and for at least one left branch (a pre-dominant) leading into the phrase's cadence. Finally, there is a third overarching principle, that of parallelism: parallel passages preferably have parallel structures. *GTTM* uses this principle in all of its theoretical components.

The principles of the shortest path, prolongational good form, and parallelism reinforce one another in the Bach phrase, but in other passages they might conflict. Although the procedures involving the shortest path are algorithmic, their interaction with prolongational good form is not fully specified; and the principle of parallelism is notoriously difficult to quantify. Hence there is a degree of flexibility in the assignment of prolongational structure.

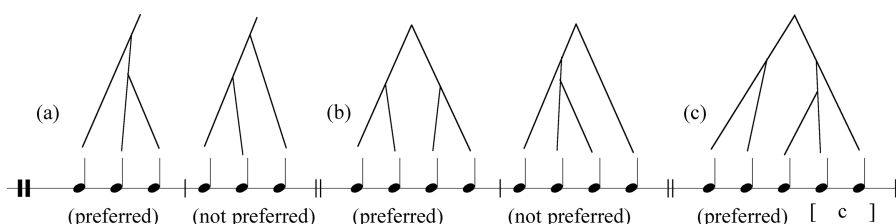


FIGURE 8. Prolongational good form: (a) recursion constraint; (b) balance constraint; (c) normative structure.

The chord distance rule calculates not only the distance between two chords x and y but also the tonal tension between them. Tension can be computed both sequentially and hierarchically. Sequential tension is measured simply from one event to the next, as if the listener had no memory or expectation. Hierarchical tension proceeds through the prolongational analysis from global to local levels in the tree structure. It is an empirical question how much listeners hear tension sequentially and how much hierarchically. No doubt they hear from one event to the next, but if listening were only sequential there would be little larger-scale coherence to the musical experience.

We turn now to the third component of tonal tension, surface dissonance. Nonharmonic tones (tones not belonging to a sounding triad) are less stable, hence tenser, than harmonic tones. Even when all the sounding tones are harmonic, the triad is more stable if it is in root position than if it is in inversion; and, to a lesser extent, it is more stable if its melodic note is on the root of the triad than if it is on the third or fifth scale degree. These factors are registered, categorically and approximately, in the surface tension rule in Figure 9. They are only approximate because tones within a category in fact differ in their degree of perceived dissonance, depending on intervallic structure, metrical position, duration, loudness, timbre, and textural location. An alternative method would be to quantify surface tension according to an established measure of sensory dissonance in the psychoacoustic literature (for instance, Hutchinson & Knopoff, 1978). This method would give rise to a continuous measure of surface tension.

However, surface tension is perceived categorically to a considerable extent. For example, in a diatonic 7-6 suspension chain, all the sevenths, major or minor, sound more or less equally dissonant. Here we take the categorical approach.

The chord distance rule and the surface tension rule combine in two possible ways to yield an overall tension value for a given event. The simpler way, stated in Figure 10a, is sequential: calculate the pitch-space distance from one event to the next and add the value for surface tension. The more complex way in Figure 10b is hierarchical: calculate the pitch-space distance from the immediately dominating event and add the value for surface tension; then add hierarchical values as inherited down the prolongational tree.

As illustration, consider Figure 11, the Grail theme from Wagner's *Parsifal*. (This is also known as the "Dresden Amen" and is familiar as such in some Protestant services. Here the theme is transposed from $A\flat$, its characteristic key, to $E\flat$ so that it can be directly compared later on to its chromatic version in $E\flat$.) The theoretically preferred analysis, following the recursion constraint and parallelism for the first four events, says that the music tenses away from the opening I until the pre-dominant ii (Event 5) in bar 2. After an elaboration of ii, the progression relaxes, in observance of normative structure, into the closing I, which repeats the opening I an octave higher. The dashed branch to Event 5 signifies an alternative branching that continues to follow the parallelism of the harmonic sequence but that removes the pre-dominant left branch required by normative structure. We shall return to this point.

Surface tension rule: $T_{\text{diss}}(y) = \text{scale degree} + \text{inversion} + \text{non-harmonic tones}$ (summed over all the pitch classes in y 's span), where

scale degree = 1 if $\hat{3}$ or $\hat{5}$ in the melodic voice, 0 otherwise;

inversion = 2 if $\hat{3}$ or $\hat{5}$ in the bass, 0 otherwise;

non-harmonic tone = 3 if a pitch class is a diatonic non-chord tone, 4 if it is a chromatic non-chord tone, 0 otherwise.

FIGURE 9. The rule for calculating surface tension.

(a)

Sequential tension rule: $T_{\text{seq}}(y) = \delta(x_{\text{prec}} \rightarrow y) + T_{\text{diss}}(y)$, where y = the target chord, x_{prec} = the chord that immediately precedes y in the sequence, $T_{\text{seq}}(y)$ = the tension associated with y , and $\delta(x_{\text{prec}} \rightarrow y)$ = the distance from x_{prec} to y .

(b)

Hierarchical tension rule: $T_{\text{loc}}(y) = \delta(x_{\text{dom}} \rightarrow y) + T_{\text{diss}}(y)$; and $T_{\text{glob}}(y) = T_{\text{loc}}(y) + T_{\text{inh}}(x_{\text{dom}})$, where y = the target chord, x_{dom} = the chord that directly dominates y in the prolongational tree; $T_{\text{loc}}(y)$ = the local tension associated with y ; $\delta(x_{\text{dom}} \rightarrow y)$ = the distance from x_{dom} to y ($= i + j + k$); $T_{\text{glob}}(y)$ = the global tension associated with y ; and $T_{\text{inh}}(x_{\text{dom}})$ = the sum of distance values inherited by y from chords that dominate x_{dom} .

FIGURE 10. Tension rules: (a) sequential tension plus surface dissonance; (b) hierarchical tension plus surface dissonance.

Included in Figure 11 are numerical values from the application of the rules in Figures 9 and 10. The first row of numbers between the staves lists surface dissonance values. The second row lists sequential tension values, obtained by calculating δ from one chord to the next and adding surface distance values. For example, the sequential distance from Event 2 to Event 3 is 7, and the surface dissonance value for Event 3 is 1; so the sequential tension associated with Event 3 is $7 + 1 = 8$. The third row similarly lists hierarchical tension values, obtained globally by adding the distance numbers next to the branches of the tree and then adding the surface distance values. Thus the hierarchical distance from Event 2 to Event 4 is $0 + 7 + 7 = 14$, and the surface dissonance value for Event 4 is 1; hence the hierarchical tension associated with Event 4 is $14 + 1 = 15$.

The same calculations appear in the tabular format in Figure 12. The events for T_{seq} in Figure 12a are listed in sequential order. The table decomposes the surface-dissonance and pitch-space factors into their component parts. The values in each row are summed to reach the total sequential tension for each event. In T_{hier} in Figure 12b, the target chords (those to the right of the arrows) are still listed in sequential order, but the source chords (those to the left of the arrows) are now listed by the immediately dominating events in the tree. For example, the notation $T_{\text{hier}}(4 \rightarrow 3)$ indicates that Event 3, because it branches from Event 4, derives its tension value from Event 4. In accordance with the hierarchical tension rule, Figure 12b includes the additional columns of “local total” and “inherited value.” The hierarchical tension for each event, given in the “global

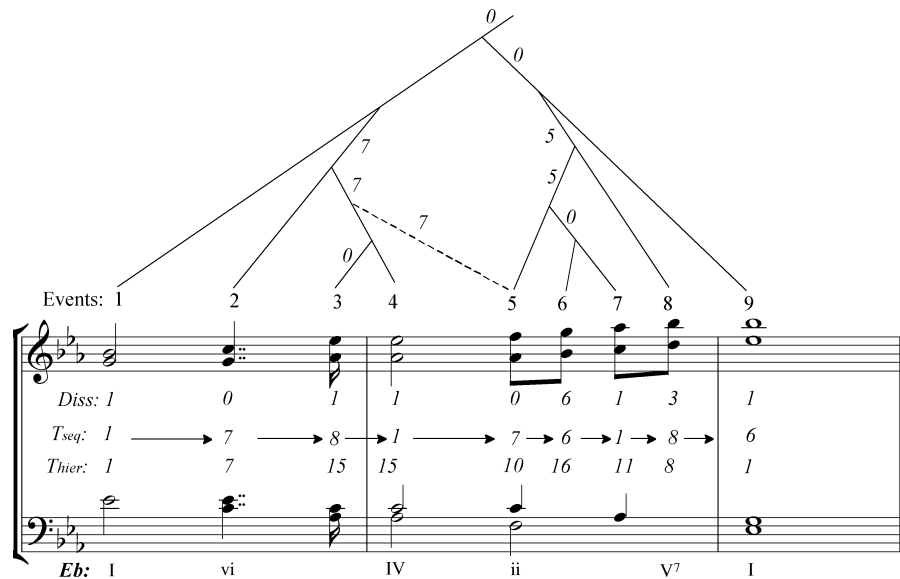


FIGURE 11. Grail theme (diatonic version) from Wagner's *Parsifal*, together with its' theoretically preferred prolongational analysis, surface dissonance values, sequential tension values, and hierarchical tension values.

(a)	<i>surface dissonance</i>			<i>pitch-space distance</i>			<i>total</i>
	<i>sc.dg.</i>	<i>inv.</i>	<i>nh.t.</i>	<i>i</i>	<i>j</i>	<i>k</i>	
T _{seq} (0→1) [I]	1	0	0	0	0	0	1
T _{seq} (1→2) [vi]	0	0	0	0	3	4	7
T _{seq} (2→3) [IV]	1	0	0	0	3	4	8
T _{seq} (3→4)	1	0	0	0	0	0	1
T _{seq} (4→5) [ii]	0	0	0	0	3	4	7
T _{seq} (5→6)	0	0	6	0	0	0	6
T _{seq} (6→7)	1	0	0	0	0	0	1
T _{seq} (7→8) [V ⁴ ₃]	0	2	1	0	1	4	8
T _{seq} (8→9) [I]	1	0	0	0	1	4	6

(b)	<i>surface dissonance</i>			<i>pitch-space distance</i>			<i>local total</i>	<i>inh. value</i>	<i>global total</i>
	<i>sc.dg.</i>	<i>inv.</i>	<i>nh.t.</i>	<i>i</i>	<i>j</i>	<i>k</i>			
T _{hier} (0→1) [I]	1	0	0	0	0	0	1	0	1
T _{hier} (1→2) [vi]	0	0	0	0	3	4	7	0	7
T _{hier} (4→3)	1	0	0	0	0	0	1	7 + 7	15
T _{hier} (2→4) [IV]	1	0	0	0	3	4	8	7	15
T _{hier} (8→5) [ii]	0	0	0	0	1	4	5	5	10
T _{hier} (7→6)	0	0	6	0	0	0	6	5 + 5	16
T _{hier} (5→7)	1	0	0	0	0	0	1	5 + 5	11
T _{hier} (9→8) [V ⁴ ₃]	0	2	1	0	1	4	8	0	8
T _{hier} (1→9) [I]	1	0	0	0	0	0	0	0	1

FIGURE 12. Tension tables for Figure 11: (a) sequential tension; (b) hierarchical tension.

total” column, equals the local total plus the inherited value.

The fourth component of the tension model is the factor of attraction. That pitches tend strongly or weakly toward other pitches has long been recognized in music theory (see *TPS*, pp. 166–167 and 188–192). Bharucha (1984, 1996) provides a psychological account of this phenomenon through the notion of anchoring, which is the urge for a less stable pitch to resolve on a subsequent, proximate, and more stable pitch. This corresponds to the account offered by Krumhansl (1979) for the effect of temporal order on tone similarity judgments. Bharucha and Larson (1994, 2004) also equate the attractive urge with melodic expectancy (Meyer, 1956; Narmour, 1990). The *TPS* attraction model extends Bharucha’s approach to include the attraction of any pitch to any other pitch and to harmonic progression. It also quantifies the relevant variables and places them within a larger cognitive theory.

Figure 13a repeats the basic space with the fifths level (level b in Figure 3) omitted, in order to make attractions

to the third and fifth scale degrees equal. Each level of the space is assigned an anchoring strength in inverse relation to its depth of embedding. Figure 13b gives the melodic attraction rule. The two factors in the equation, combined by multiplication, are the ratio of anchoring strengths of two pitches and the inverse square of the semitone distance between them. The distance factor is estimated to behave as in Newton’s classical gravitational equation. The inverse-square factor renders miniscule attractions between pitches that are more than a major second apart. To convey the behavior of the rule, Figure 13c lists a few attractions to diatonic neighbors in the context I/C. The pitch B is highly attracted to C because the two pitches are a semitone apart and C is more stable. D is less attracted to C because it is two semitones away. F is more attracted to E than E is to F because of their inverse anchoring strengths.

The attraction rule applies not only to individual lines but also to each voice in a progression. As stated in the harmonic attraction rule in Figure 14, these values

(a)

<i>Anchoring strength</i>	<i>The basic space with the fifths level omitted</i>											
4	0											
3	0		4		7							
2	0		2		4 5		7		9		11	
1	0		1 2 3		4 5		6 7		8 9		10 11	

(b)

Melodic attraction: $\alpha(p_1 \rightarrow p_2) = s_2/s_1 \times 1/n^2$, where p_1 and p_2 are pitches, with $p_1 \neq p_2$; $\alpha(p_1 \rightarrow p_2)$ = the attraction of p_1 to p_2 ; s_1 = the anchoring strength of p_1 and s_2 = the anchoring strength of p_2 in the current configuration of the basic space; and n = the number of semitone intervals between p_1 and p_2 .

(c)

$$\begin{aligned}\alpha(B \rightarrow C) &= 4/2 \times 1/1^2 = 4/2 = 2 \\ \alpha(D \rightarrow C) &= 4/2 \times 1/2^2 = 4/8 = 0.5 \\ \alpha(F \rightarrow E) &= 3/2 \times 1/1^2 = 3/2 = 1.5 \\ \alpha(E \rightarrow F) &= 2/3 \times 1/1^2 = 2/3 = 0.67\end{aligned}$$

FIGURE 13. Melodic attractions: (a) The basic space minus the fifth level and with anchoring strengths indicated by level; (b) the melodic attraction rule; (c) some computed attractions between scalar adjacencies in the context I/C.

are summed and then divided by the value for the chord distance rule to obtain the overall attraction value from one chord to the next.

Figure 15 applies the harmonic attraction rule to the first and last progressions in Wagner's Grail theme. Where a pitch repeats, "null" is designated because the attraction rule does not apply to repeated pitches. The values of α are summed to the combined realized voice-leading value (α_{rvl}), which is divided by the δ value to give the final realized harmonic attraction value (α_{rh}). Notice the extreme differences between the α_{rh} values for Prog(1→2) and Prog(8→9). In the former, the progression I→vi is only moderately strong and includes repeated notes; in the latter, the progression V⁷→I is very strong and resolves by half step in two voices. Indeed, the strongest harmonic attraction is from a dominant seventh chord to its tonic, because of the powerful attractions of the leading tone to the tonic

and the fourth to the third scale degree and because of the short distance from the dominant to the tonic chord. This is why (aside from statistical frequency) the expectancy for a tonic chord is so high after a dominant-seventh chord.

Attractions in *TPS* are computed not only from event to event at the musical surface but also from event to event at immediately underlying levels of prolongational reduction. The resulting sets of numbers, however, are not integrated into a single attraction measure across reductional levels. Depending in part on tempo, underlying levels presumably contribute to the overall result in increasingly smaller amounts as the analysis abstracts away from the surface. (Margulis, 2005, proposes a mechanism for this step.) In this study we dispense with underlying levels of attraction.

Figure 16 shows the surface attraction values for the Grail theme. The numbers appear between events because

Harmonic attraction rule: $\alpha_{rh}(C_1 \rightarrow C_2) = c[\alpha_{rvl}(C_1 \rightarrow C_2)/\delta(C_1 \rightarrow C_2)]$, where $\alpha_{rh}(C_1 \rightarrow C_2)$ is the realized harmonic attraction of C_1 to C_2 ; constant $c = 10$; $\alpha_{rvl}(C_1 \rightarrow C_2)$ is the sum of the realized voice-leading attractions for all of the voices in C_1 ; and $\delta(C_1 \rightarrow C_2)$ is the distance from C_1 to C_2 , with $C_1 \neq C_2$.

FIGURE 14. Harmonic attraction rule.

<p>(a)</p> <p><i>Prog</i>(1→2)</p> $\alpha(B\flat \rightarrow C) = 0.17$ $\alpha(G \rightarrow G) = \text{null}$ $\alpha(E\flat \rightarrow E\flat) = \text{null}$ $\alpha(E\flat \rightarrow C) = 0.06$ $\alpha_{\text{rvl}}(I/E\flat \rightarrow vi/E\flat) = 0.23$ $\alpha_{\text{rh}}(I/E\flat \rightarrow vi/E\flat) = 10 \times (.23/7) = 0.33$	<p>(b)</p> <p><i>Prog</i>(8→9)</p> $\alpha(B\flat \rightarrow B\flat) = \text{null}$ $\alpha(D \rightarrow E\flat) = 2.0$ $\alpha(A\flat \rightarrow G) = 1.5$ $\alpha(F \rightarrow E\flat) = 0.5$ $\alpha_{\text{rvl}}(V^7/E\flat \rightarrow I/E\flat) = 4.0$ $\alpha_{\text{rh}}(V^7/E\flat \rightarrow I/E\flat) = 10 \times (4/5) = 8.0$
---	--

FIGURE 15. Two applications of the harmonic attraction rule.

they apply to relations between events. Where the harmony does not change, as in events 3-4 and 5-7, a single attraction value obtains.

There is a complementary relationship between tension and attraction numbers. Where the music tenses away from the tonic, attractions are realized on less stable pitches and chords. Hence where tension numbers rise, attraction values tend to be small. But where the music relaxes toward the tonic, attractions are realized on more stable pitches and chords; tension numbers decline and attraction values rise. A high attraction value in effect constitutes a second kind of tension—not the tension of motion away from stability but the tension of expectation that the attractor pitch or chord will arrive.

A further general point about tension and attraction concerns numerical quantification. As Klumpenhouwer (2005) points out, the theory's numbers measure different entities in the different components: in the dissonance component, chord inversions and nonharmonic tones; in the distance model, steps on cycles of fifths and noncommon tones; in the attraction component, pitch stabilities and distances. As numerical values, then, these might be considered incommensurate (for example, a 2 for inversion in the dissonance component is not exactly the same as a 2 for the k distance between chords). One approach to this issue would be to find coefficients for the different variables to express the relative strength of their units of measurement. We have found, however, that coefficients are not needed for the tension rules; that is, the numbers already express the

relative strength of the variables in question. However, the attraction rules yield incommensurable output numbers compared to those of the tension rules. Empirical data suggest that coefficients are needed when tension combines with attraction. For this, we take a practical rather than theoretical solution through the mathematical technique of multiple regression, which weights the two sets of numbers to find the best fit between the tension and attraction curves.

Before proceeding, it should be noted that the melodic attraction rule (Figure 13b) stands on weaker empirical grounds than does the chord distance rule (Figure 4). Experimental results guided the development of the distance rule. (However, the output of the elaborated form of δ , the chord/region distance rule Δ [TPS, p. 70], which employs the pivot-region concept, proves to be empirically less successful, and we shall not invoke it.) The attraction rule, in contrast, was developed by a blend of theoretical and intuitive considerations without much supporting empirical data. Several aspects of the rule can be criticized. First, it is unclear that a multiplicative rather than additive relationship should obtain between the stability (s_2/s_1) and proximity ($1/n^2$) parts of the equation. Second, as Larson (2004) and Samplaski (2005) observe, there is arbitrariness in the reduction of five levels of the basic space (Figure 3) to four when calculating attractions (Figure 13a). Third, the inverse factor for proximity eliminates the attraction of a pitch to itself because of the impossibility of a zero denominator. Pitch repetition may indeed be a case where intuitions of attraction and expectation

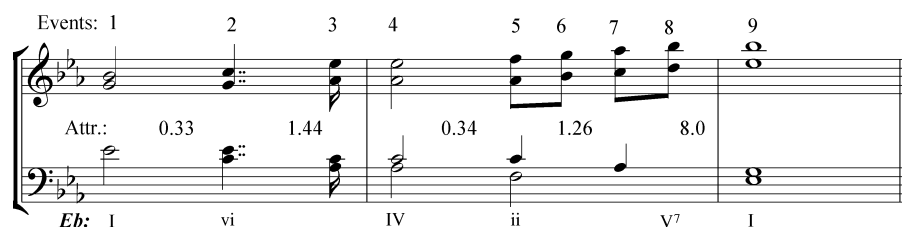


FIGURE 16. Attraction values for the Grail theme.

diverge. One may expect a pitch to repeat, but it seems more natural to think of a pitch as being attracted only to other pitches. Nevertheless, the exclusion of pitch repetition from the calculations leaves a gap in the theory. Fourth, the specific form of the inverse factor, $\frac{1}{n^2}$, appears to create too steep a curve; that is, the drop from great attraction at the half-step distance to less attraction at the whole-step distance to very little attraction at the minor-third distance seems too extreme (Margulis, 2005). The obvious alternative, $\frac{1}{n}$, yields too flat a curve. An intermediate curve is possible, but the theoretical and empirical bases for such a solution are unclear. Fifth, and perhaps most importantly from a theoretical perspective, the measurement of proximity only by semitones may be too simple a metric. Larson (2004) cites evidence in Povel (1996) that stepwise arpeggiated intervals—that is, between adjacent members of a triad or between the fifth and the tonic—yield greater attractions than predicted by $\frac{1}{n^2}$ or $\frac{1}{n}$. Krumhansl (1979) and *CFMP* (Table 5.1) also find high relatedness ratings for triad members. This evidence fits the discussion in Chapter 2 of *TPS* about pitch proximity, step motion, and linear completion. It appears that that discussion, in which stepwise motion is seen as pertaining to the alphabet in question at a given level of the basic space, should have informed the formulation of pitch attractions in Chapter 4.

Despite these reservations, the principles behind the attraction rule, stability and proximity, remain the central factors in a treatment of melodic attraction. We have tried the alternatives of five instead of four stability levels and of proximity by $\frac{1}{n}$ instead of $\frac{1}{n^2}$, but the resulting values do not lead to improvements over those of the original rule with respect to the empirical data. Nor are there enough instances of voice-leading arpeggiation in our examples to force a stratified treatment of melodic proximity. Our project is to test the success or failure of the *TPS* theory of tension and attraction, and we leave theoretical refinements of the attraction rule for future research.

Experimental Approach

The participants in the experiments under discussion were musically trained students at Cornell University with relatively little training in music theory compared to the extent of their instruction on musical instruments or voice. (More details of music backgrounds and other details of the experimental method can be found in Appendix A.) They were tested for tension responses for Wagner's Grail theme from *Parsifal* in its diatonic and chromatic versions, a Bach chorale, a

chromatic Chopin prelude, and a passage from Messiaen's *Quartet for the End of Time*. The data were compared to the model's predictions. (A Mozart sonata movement that received a similar treatment is not discussed in this paper; see Krumhansl, 1996, and Lerdahl, 1996.)

The tests for the Wagner and Bach excerpts were conducted in two ways, the stop-tension task and the continuous-tension task. In the stop-tension task, the first event was sounded, at which point the participants rated its degree of tension; then the first and second events were sounded and the participants rated the tension of the second event; then the first, second, and third events were sounded, and so on, until the tension associated with each successive event was recorded. In the continuous-tension task, which was done for all excerpts, the participants interacted with a graphic interface that enabled them to move a slider right and left on the computer screen using a mouse, in correspondence with their ongoing experience of increasing and decreasing tension. The advantage of the stop-tension task is that it records the response precisely for the event that is evaluated. Its disadvantage is that it is rather artificial and prohibitively time-consuming for long excerpts. The advantage of the continuous-tension task is that it encourages a spontaneous response to intuitions of tension in real time. Its disadvantage is that there is a lag time, for which an approximated correction must be made, between the sounded events and the physical response of moving the mouse. Perhaps surprisingly, the results from the two tasks yielded almost the same results for the short passages where both tasks were used. For the longer Chopin and Messiaen selections, however, it was practical to employ only the continuous-tension task. The participants in the study by Krumhansl (1996) using this method varied in the extent of their musical training, but training had little effect on the tension judgments.

As mentioned, the analyses combine tension and attraction values to achieve an overall measure of tension. We follow three conventions in this respect. First, even though an attraction number does not adhere to a single pitch but represents a relation between two successive pitches x and y , we assign the number to x , in effect claiming that it is at x that the experience of attraction most saliently takes place. In this way, each event has two numbers associated with it, one for δ and the other for α . Second, the harmonic attraction rule (Figure 14) has δ in the denominator and hence requires that $\delta \neq 0$. This creates a problem when the voice leading moves but the harmony does not progress. In such cases, we repeat the value for δ from the point at which the harmony last changed (as in Figure 16, Events 3-4 and 5-7). Third, a

virtual attraction can be computed from Event x to any possible Event y , and, in particular, to the y with the highest attraction value. Instead we calculate only the realized attraction, that is, from x to the y that actually follows. It might be argued, especially for the stop-tension task, that the strongest virtual attraction, when it is not the same as the realized attraction, should be calculated, on the view that the strongest attraction corresponds to the strongest expectation. Expectations, however, depend not only on strongest attractions but also on schematic patterns that lie beyond current formalization. To calculate to an event that does not occur would be somewhat speculative in this context. It suffices as a first approximation to rely on the definiteness of realized attractions.

A larger methodological point concerns the interaction between prediction and data. It is sometimes thought that an experiment simply tests a preexisting theory. Yet experimental data can give rise to a theory; this in fact was the case for the construction of the pitch-space model. In a healthy science, it often happens that a fruitful exchange develops between theory and experiment. Such is the case here. If the data suggest that the predictions are faulty, principled ways are sought within the model to reach predictions that achieve a better empirical result. These reevaluations are principled in the sense that they are constrained by the general assumptions and specific formalisms of the theory. This process can go back and forth a number of times. One must of course be careful not to adjust the theory simply in order to fit the data. Rather, the data can illuminate how listeners construe tension, suggesting interpretations within the model that are both theoretically acceptable and more predictive. In this way the theory can be improved. Furthermore, in our view it is not enough to achieve a statistically significant overall correlation. What is wanted, in addition, is an explanatory account of why the model succeeds or fails at any given point in the analysis.

In this back-and-forth process there are two kinds of flexibility within the theory. First, sequential or hierarchical tension can be computed, each with or without attractions. Second, unlike the tension and attraction rules (all those that incorporate δ and α), which are algorithmic, the derivation of prolongational structure involves gradient preference rules, which interact with one another in search of an optimal solution (see the discussions in *GTTM* and *TPS*; also Temperley, 2001).

Preferential conditions arise in three ways. First is the interaction between the principles of the shortest path, prolongational good form, and parallelism. Second, when there is a shift from a right- to a left-branching

pattern, the event where the shift takes place can attach either way, depending on the shortest path and good form. Third, it is not always clear where to locate an event in pitch space; that is, there can be ambiguity about the identity of a chord or the exact moment of a modulation. As a result of these factors, a passage of music yields not a single prolongational analysis but a limited range of preferred analyses. The data can point in any given case toward which theoretically viable prolongational analysis conforms best to listeners' responses.

Analyses in a Diatonic Framework

Wagner Theme, Diatonic Version

We begin with the diatonic Grail theme from Wagner's *Parsifal*, shown in Figure 11. Figure 17 records the nine events of the excerpt on the x-axis and tension responses from the stop-tension task on the y-axis. The dashed line represents the sequential tension values from Figure 12a, without the inclusion of attraction values, and the solid line shows the data from the averaged listeners' responses. The fit is quite poor: $R^2(1,7) = .08$, $p = .46$, $R^2_{adj} = -.049$.

Some words of explanation may be helpful. For each correlation, we present the following information about the statistical test. The first number R^2 is the proportion of variation in the data that is accounted for by the model. It is associated with two numbers, the degrees of freedom. The first degree of freedom indicates the number of predictor variables in the model. In this case there is one variable, sequential tension. The second degree of freedom is the number of data points (in this

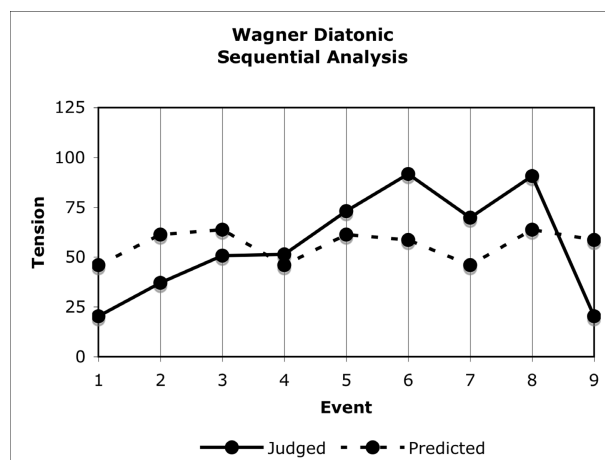


FIGURE 17. Sequential tension analysis of the Grail theme from *Parsifal*.

case 9, one for each event in the music) minus two. The subtraction of two results from the regression “using up” two degrees of freedom for the parameters it determines. The regression finds the best-fitting linear model predicting the data from the variable(s). The model determines the optimal values for the slope of the line (one parameter) and the intercept of the line (the second parameter). Hence the number two is subtracted from the number of data points going into the regression to give the second degree of freedom. If the model can find a perfect fit between the data and the variable(s), R^2 would equal 1.0. In general, the value is less than this, and its significance is measured by the probability, p , which is the next number given in the statistical report. By convention, a statistic, such as R^2 , is considered significant when the p value is less than .05. The probability depends on both the size of R^2 and the degrees of freedom. The last number given is the adjusted R^2 , R^2_{adj} . The R^2_{adj} is the R^2 value adjusted to make it more comparable with other models for the same data that have different numbers of degrees of freedom.

Methods such as time-series analysis or functional data analysis are not appropriate here. Our objective is to determine whether the judgments fit the quantitative predictions of the model for each musical event. For this we need a single number for judged tension for each event.

The conclusion from this statistical test for Figure 17 is that listeners do not hear this passage in a simple sequential manner. The R^2 value is only .08, which means that the sequential tension variable accounts for only 8% of the variability in the tension judgments, and the p value of .46 tells us that this is an unimpressive result. Graphically, this is apparent in Figure 17 where the two lines do not follow each other closely.

The second analysis is another single variable model, using the attraction values displayed in Figure 18. These are the attraction values from Figure 16, without the inclusion of tension values, against the listeners' responses. The fit is improved but still not good: $R^2(1,7) = .35$, $p = .09$, $R^2_{adj} = .26$.

Figure 19 combines Figures 17 and 18 by adding the attraction values to the sequential tension values. Multiple regression weights the two sets of numbers to achieve a best-fit solution, and assigns a probability to each of the predictor variables. These will be denoted $p(\text{attraction})$ for the probability of attraction and $p(\text{tension})$ for the probability of the total tension predictor. Each of these is shown with a standardized beta value, β . The β weights are the coefficients in the linear model predicting the data from the predictor variable

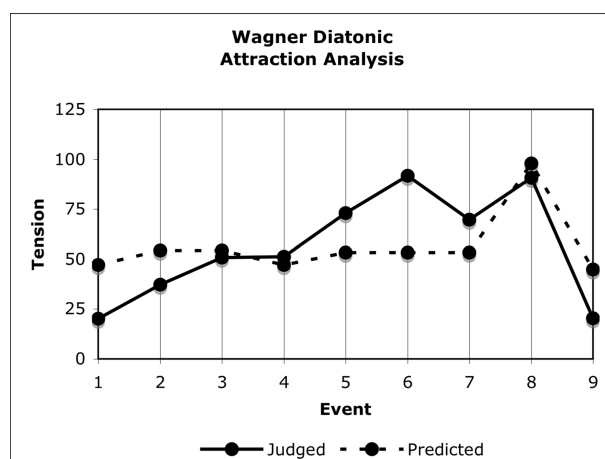


FIGURE 18. Attraction analysis of the Grail theme.

(after they have been standardized to have the same mean and standard deviation). The picture is better than in Figure 17 but no better than in Figure 18: $R^2(2,6) = .35$, $p = .28$, $R^2_{adj} = .13$; $p(\text{attraction}) = .17$, $\beta = .58$; $p(\text{tension}) = .96$, $\beta = .02$. The higher p value for R^2 is the penalty for using two predictor variables rather than one, thus increasing the first degree of freedom. Or, to put it another way, when attraction is included in the model, adding the sequential tension values does not improve the fit (the p value for sequential tension in the multiple regression is .96, which means that adding it has virtually no effect). This analysis confirms the conclusion that the strict sequential treatment of tension does not contribute to the fit of the data.

Let us abandon the sequential-tension approach and consider instead the theoretically preferred prolongational

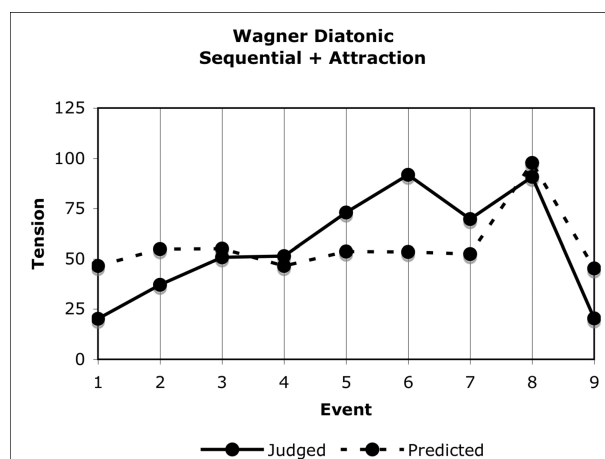


FIGURE 19. Combined sequential + attraction analysis of the Grail theme.

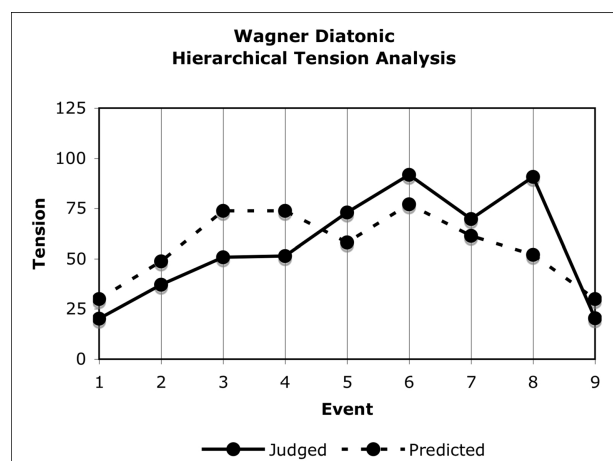


FIGURE 20. Tension graph for the theoretically preferred hierarchical analysis of the Grail theme.

analysis in Figure 11 together with its derived hierarchical tension analysis in Figure 12b. At first we ignore attraction values. The resultant graph in Figure 20 achieves a better correlation than the previous analyses: $R^2(1,7) = .43$, $p = .056$, $R^2_{adj} = .35$. However, the predicted values are too high for Events 1-4 and too low for Events 5-8.

Figure 21 adds the attraction values in Figure 16 to the tension values in Figure 12b. Now the correlation is quite good and statistically significant: $R^2(2,6) = .75$, $p = .016$, $R^2_{adj} = .66$; $p(\text{attraction}) = .03$, $\beta = .56$; $p(\text{tension}) = .02$, $\beta = .63$. The most notable change in Figure 21 compared to Figure 20 is the raising of the predicted curve at Event 8 (V^7). In Figure 20 the tension model

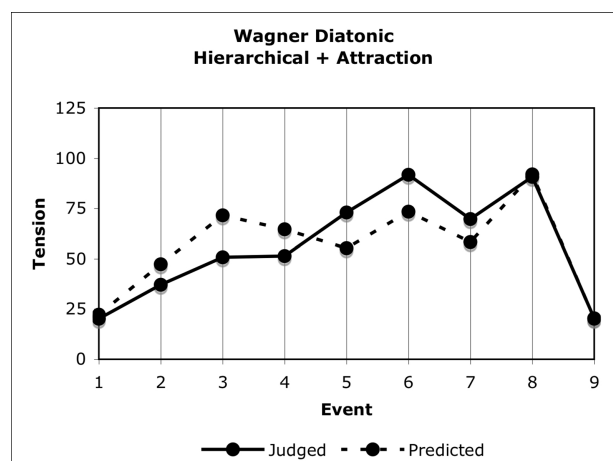


FIGURE 21. Combined hierarchical (theoretically preferred) + attraction analysis of the Grail theme.

correctly assigns relaxation into the cadence, but participants experience greater tension at the V^7 chord than shown there. This happens because the V^7 is highly attracted to the following tonic resolution, an effect realized in Figure 21 by the inclusion of attraction values. Discrepancies remain, however. The predictions for Events 3-4 are still too high and those for Events 5-7 are too low.

These shortcomings can be overcome through a revision of the prolongational analysis. In the original analysis in Figure 11, there is equilibrium between right and left branching (following the balance constraint), with Event 5 (the ii chord) interpreted as a pre-dominant to the cadence (following normative structure). The analysis in effect claims that, beginning at Event 5, the listener already expects the resolution on Event 9. But it is harder to anticipate prospectively than it is to remember retrospectively. Besides, Event 5 continues from the previous events the harmonic sequence of descending thirds with a rising melodic second. It is easier to hear instead the analysis in Figure 22, in which the principle of parallelism wins over those of branching balance and pre-dominant function. The only difference is that Event 5 is now a right instead of left branch; Events 6 and 7 attach to Event 5 as before. This single change leads to alterations in tension values for Events 5-7, as listed between the staves. In this interpretation, the tension of the harmonic sequence continues through the elaboration of ii in Events 6-7 and is released only at the cadence in Events 8-9. Attractions remain as before. The result is the almost perfectly matching curves in Figure 23: $R^2(2,6) = .97$, $p < .0001$, $R^2_{adj} = .97$; $p(\text{attraction}) < .0001$, $\beta = .58$; $p(\text{tension}) < .0001$, $\beta = .79$.

Three broad conclusions can be drawn from this analysis of the Grail theme. First, attractions must be incorporated into the predictions. Second, listeners hear tension hierarchically more than sequentially. Third, unless schematic intuitions are strong, listeners tend to construe events in a right-branching manner, that is, in terms of previous rather than following events.

Are the stop-tension data related to the continuous-stop data? Krumhansl (1996) found that the discrete predictions of the TPS model could provide a good fit to the continuous tension judgments by assuming an integration time of 2.5 seconds. In the present case, this approach is adapted to ask whether the continuous-tension data could be predicted by the stop-tension data, assuming the same integration time. The calculation assumes that the values of past events are degraded as an inverse exponential function with a half-life of 0.5 seconds. The continuous data are plotted as a solid

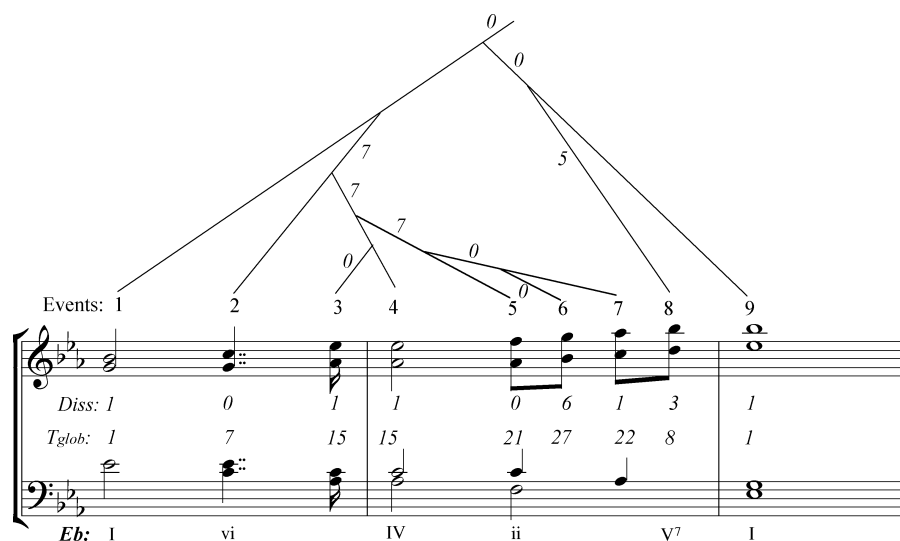


FIGURE 22. Prolongational analysis of the Grail theme, with Event 5 reinterpreted as right branching.

line in Figure 24 together with the values calculated from the stop-tension data. A high degree of agreement is reached: $R^2(1,104) = .95$, $p < .0001$, $R^2_{adj} = .95$. This is of interest because the participants performed the stop-tension task before the continuous-tension task. This means that when they performed the stop-tension task, they had not heard the music beyond the chord that they were judging. The extent to which the two tasks converge suggests that listeners were responding to the sounded events rather than to events they anticipated because of memory from previous listening. Although the analyses will not be presented here, the stop-tension and continuous-tension data are similarly related for

the other two excerpts for which they are available (the chromatic version of the Wagner Grail theme and the Bach chorale).

Bach Chorale

On the basis of the discussion of the Wagner excerpt, the remaining analyses follow the hierarchical rather than sequential tension model and incorporate attractions as part of the overall prediction of tension. We first consider the Bach chorale “Christus, der ist mein Leben.” Its prolongational analysis is divided, for reasons of space, between Figure 25 and Figure 26. The top

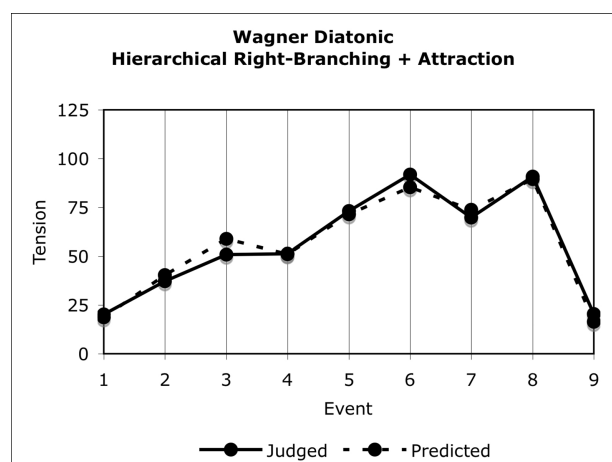


FIGURE 23. Combined hierarchical (right-branching) + attraction analysis of the Grail theme.

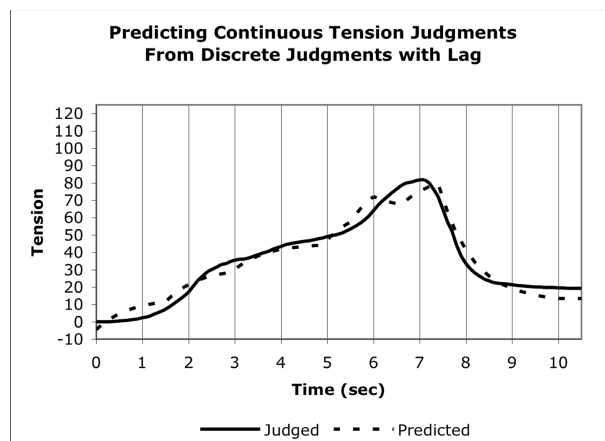


FIGURE 24. Comparison of the continuous-tension data (solid line) with predictions from the stop-tension data for the Grail theme, after the latter are integrated over 2.5 seconds with an exponential decay with half life 0.5 sec.

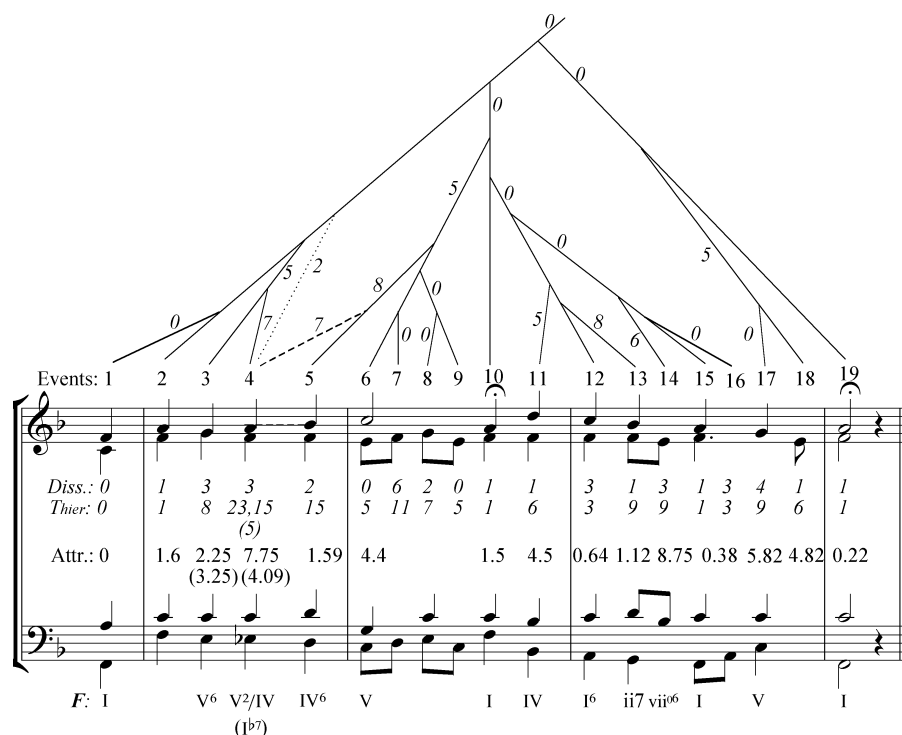


FIGURE 25. Analysis of the Bach chorale, phrases 1-2.

branches, all of which represent the tonic I/F (hence $\delta=0$ in the tree), should be understood as connecting together. Event 2 in Figure 25 attaches to Event 41 in Figure 26, and the designation for Event 19 in Figure 26 refers to Event 19 in Figure 25. The predicted values of surface dissonance, hierarchical tension, and attraction appear between the staves. (Incidentally, Event 34 branches differently than does the equivalent first event in Figure 7. Here it connects not to the final cadence [Event 41] but back to Event 19, showing the return to I³/F. This happens because a prolongational analysis always makes the most global connection possible. In Figure 7 the context was a single phrase; here it is the entire chorale.)

Figure 27 shows the fit of the empirical data with the predictions in Figures 25-26: $R^2(2,38) = .79$, $p < .0001$, $R^2_{adj} = .78$; $p(\text{attraction}) < .0001$, $\beta = .47$; $p(\text{tension}) < .0001$, $\beta = .67$. The high correlation is all the more impressive given that a correlation tends to decrease as the number of events increases (because there are more possible points of deviation, as shown in the second degree of freedom). Attraction and tension are both individually significant in the multiple regression.

The analysis in Figures 25-26 departs from the *TPS* analysis of the chorale in two places. The first concerns the interpretation of Event 4 in Figure 25. In *TPS* it is

conventionally treated as a secondary dominant, IV₂/IV, and by the shortest path attaches to the following IV. But this solution, shown by the dashed branch, gives a high tension value of 23 because of the double inheritance from IV (8 + 5). The right-branching alternative, the solid branch coming from the previous V⁶, takes a longer local path but achieves a better balance between right and left branching in the phrase as a whole, and it gives a moderate tension value of 15. Olli Väisälä (personal communication, October 26, 2004) points out, however, that the Roman numeral analysis of IV₂/IV itself violates the principle of the shortest path. The most efficient interpretation of Event 4 is instead as I/F with a flatted seventh in the bass, yielding a low tension value of 5. This option is shown in parentheses in Figure 25 and by the dotted branch in the tree. In this view, Event 4 is I³ at an immediately underlying level, transformed at the musical surface by the chromatic descent in the bass. (Imagine Event 4 with F3 instead of Eb3 in the bass; the progression makes perfect sense.) Of these solutions, the best match with the data is the intermediate one with the tension value of 15, and this is what we have followed here.

The second departure from the *TPS* analysis concerns the point at which the third phrase shifts from F to C.



FIGURE 26. Analysis of the Bach chorale, phrases 3-4.

In *TPS*, the reorientation is taken to occur on the down-beat of bar 5 with V_5^6/C , as illustrated in Figure 28a. This interpretation treats the melodic F5 on the third beat of bar 5 as a neighboring $\hat{4}$ between a prolonged $\hat{3}$ in C. The resulting tension values, however, are too high at the F5. Väisälä suggests instead the analysis in Figure 28b, in which the shift to F takes place later in the phrase. In this interpretation, which we have taken here, the F5 is not a mere neighbor in C but is the goal, $\hat{8}$ in F, of a linear progression from the C5 that begins the phrase.

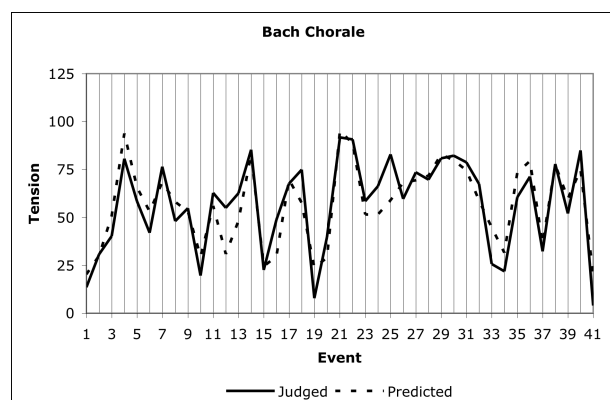


FIGURE 27. Tension graph for the Bach chorale.

This reading leads to a different prolongational tree and fits better with the data.

These alternative interpretations of the first and third phrases illustrate the gradient nature of prolongational derivational process and how empirical data can illuminate which “preferred” interpretation may best conform to listeners’ intuitions. It is noteworthy that both instances involve choices in Roman-numeral analysis. From the present perspective, Roman-numeral analysis is not just a pedagogical labeling device but is a means of establishing location in pitch space. Different spatial locations yield different distances, hierarchical relationships, and degrees of tension.

There are a few places where the model cannot find a good fit with the data. Event 12 has too low a tension value because, as I^6 prolonged from I, it inherits no tension. Yet it also acts as a passing chord in a progression of outer-voice parallel 10ths. The theory does not yet have a way of addressing this voice-leading pattern. There is also a poor fit at Events 24-25 (this would be the case also under the *TPS* interpretation in Figure 28a). These events are embellishing 16th notes of little importance to the experience of tension. However, the stop-tension task brings attention to them. The model does not yet take into account the effect of relative duration, so that these fleeting events have more weight in the statistical analysis than they ought to have.

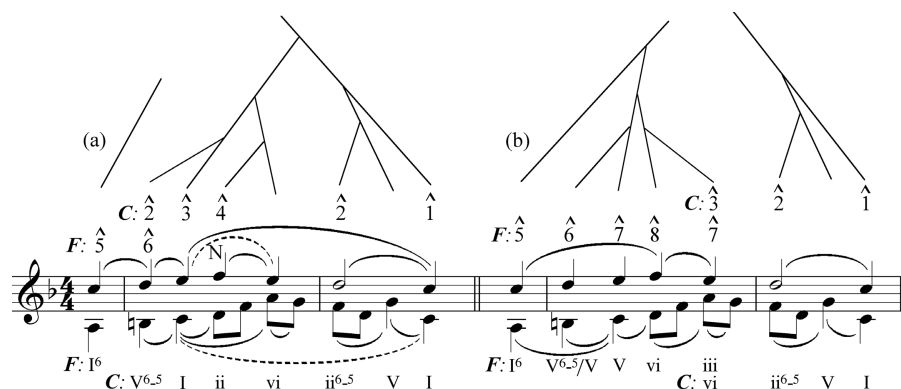


FIGURE 28. Alternative analyses of the third phrase of the Bach: (a) as in *TPS*; (b) with a delayed tonicization of C. (Only the soprano and bass lines are shown.)

Both deviations suggest directions in which the theory might be improved.

Chopin Prelude

Chopin's E major Prelude (analyzed in *TPS*, Chapter 3) is an exceptionally concentrated example of nineteenth-century chromaticism. We assume a prior reduction of the Prelude's surface to block four-part harmony. Figures 29-31 display the *TPS* prolongational analysis of the Prelude's three phrases. Each phrase begins with the same chord (I^5/E), so that, at a global level not shown, Event 17 branches off Event 1 and Event 33 off

Event 17; finally, Event 1 attaches to Event 47. As prolongations of the tonic, all these events inherit 0, and the patterns of tension and relaxation take place within the phrases.

A number of details in the figures require comment. In Figure 29, Events 6 and 8 could be regarded as separate chords (vii^{o6} and iii^6 , respectively), but it is equally valid to treat them as voice-leading anticipations to the ensuing chords (the D^\sharp in Event 6, the G^\sharp in Event 8). The latter interpretation better fits the data and is taken here. In the tree, the indication "1[0]" means that Event 16 inherits 0 from Event 12 (since both are V chords) but that the seventh (A3) in Event 16 adds 1 to its local

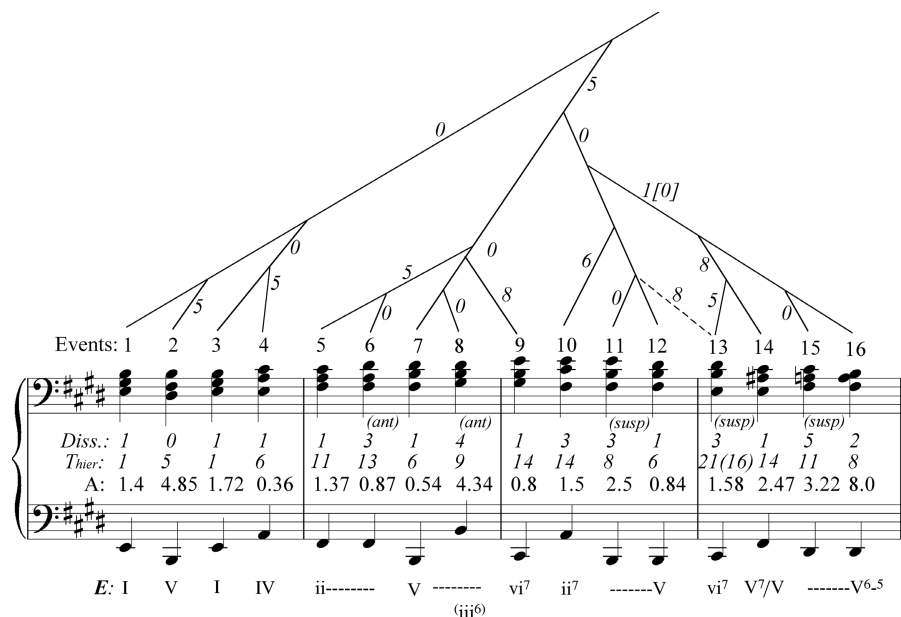


FIGURE 29. Analysis of the first phrase of Chopin's E major Prelude.

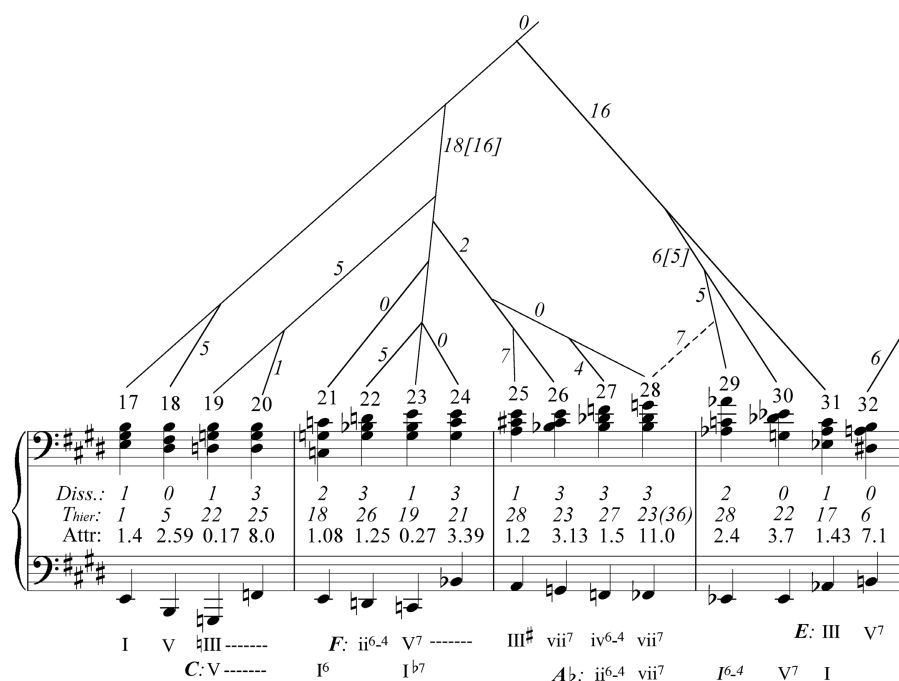


FIGURE 30. Analysis of the second phrase of Chopin's E major Prelude.

δ value. The bracketed numbers in Figure 30 to Events 23 and 30 are to be understood similarly. The dashed branch to Event 28 is an alternative interpretation that gives $T_{\text{hier}} = 36$; this result better fits the data and is adopted. Event 44 in Figure 31 also offers contrasting options. Here $T_{\text{hier}} = 6$ is too low and $T_{\text{hier}} = 45$ too high

compared to the data. We take the first option since it takes a shorter path.

Before considering the statistical fit to the data, let us note how these continuous-tension data were prepared for the analysis. The discrete values shown as "judged" in Figure 32 are the average of the listeners' tension

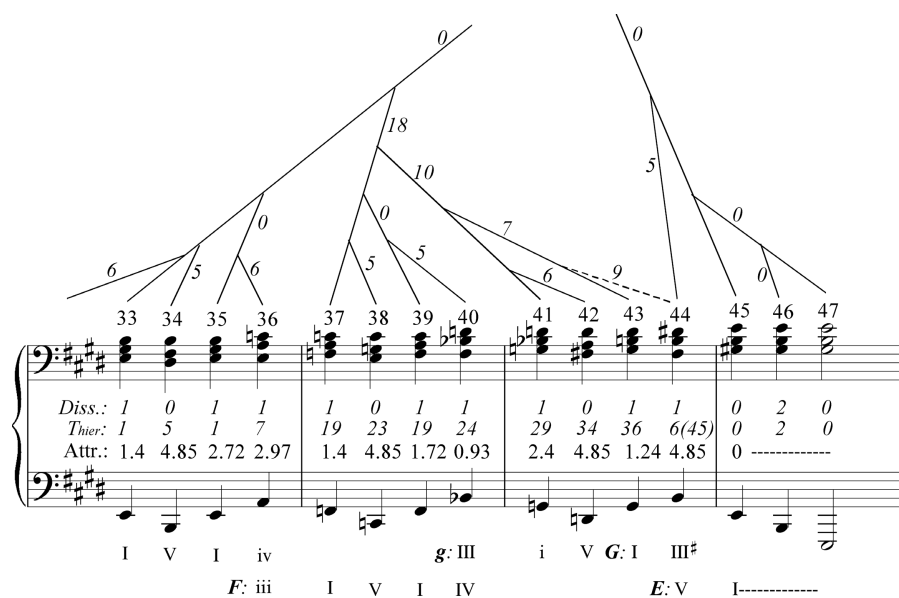


FIGURE 31. Analysis of the third phrase of Chopin's E major Prelude.

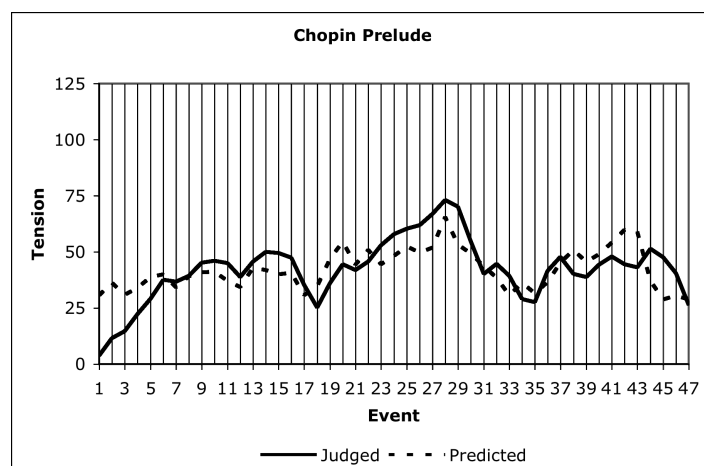


FIGURE 32. Tension graph for the *TPS* analysis of the Chopin prelude.

judgments from the onset of each event to the onset of the next event. The slow tempo (two seconds per chord) suggested that this would give a representative value for the tension of each event. The motivation for finding these discrete values was that it was desirable to work with a single number for each event as various theoretical analyses were considered. This approach makes fewer assumptions than the exponential decay model used in prior treatments of the continuous response method (Krumhansl, 1996).

Figure 32 shows the fit of the predictions in Figures 29-31 to the data: $R^2(2,44) = .42$, $p < .0001$, $R^2_{adj} = .40$; $p(\text{attraction}) = .34$, $\beta = .11$; $p(\text{tension}) < .0001$, $\beta = .62$. The correlation is not strong. Although the overall probability is low, the contribution of attraction does not approach statistical significance. We shall find a better solution, but before doing so let us review the main trouble spots. First, the discrepancy for Events 1-5 is an artifact of the continuous-tension task and can be discounted: the position of the slider was initially set at 0 and participants needed to hear a few events before they were able to position the slider near an appropriate level of overall tension. The predicted tension is too low, however, for part of the rest of the first phrase (especially Events 14-16) and much of the second phrase (Events 23-29). The fit in the third phrase is particularly poor, with the predicted values too high (Events 38-32) and then too low (Events 44-46).

The difficulty with Events 14-16 seems to be that in the prolongational analysis Event 12 inherits no tension from Event 7 (the two are identical), yet listeners pay attention instead to the slow descent of the melody. The situation is comparable to that of the second phrase of the Bach (Event 12 in Figure 25): in both cases, the linear

melodic progression maintains tension that the theory does not account for. The model's predictions for Events 23-29, in contrast, could be increased by a different analysis within the theory. The *TPS* analysis in Figure 30 follows Aldwell and Schachter (1979) by interpreting Events 24-28 as a prolongation of an enharmonically shifting diminished seventh chord; thus the A major and B \flat minor chords (Events 25 and 27) are assigned passing status. In another plausible analysis, Events 24-27 would recursively branch to the right, on the rationale that the listener is sufficiently baffled by the intense chromaticism that the only recourse is to hear each event in terms of the immediately preceding one. This tree structure would increase the predicted tension to correspond rather well to the data. We refrain from presenting this alternative only because of another option to be discussed shortly.

The third phrase presents the greatest problems, beginning with the large distance value assigned to the move to F at Event 37. As mentioned in *TPS* (p. 78), δ may obtain too great a distance between I and \flat II; Events 38-43 then inherit this value. In addition, listeners tend to lag in their responses when presented with distant modulations; they need time to adjust to the new context (see Krumhansl & Kessler, 1982, for related evidence). The data shows this in the descending curves between Events 37-39, 41-43, and 45-47. In each case, the local I-V-I progression gradually establishes the new tonic for the listener, even though in the prolongational analysis the second I is a repetition of the first. There is a clash between final-state analysis and real-time listening. The conflict is most severe at the return to E at the final cadence (Events 44-45). The listener expects a repeat of the sequential pattern in the previous bars, I-V-I-IV,



FIGURE 33. Contour values (imitating scale degrees) for the second phrase of the Chopin.

but the second I is followed instead by V/E. This startling modulation takes two more beats to process. The model has no way to register these factors.

Nevertheless, the predictions improve markedly by the injection of a new factor: melodic contour. Observe that in the first two phrases the judged tension rises and falls in waves that correspond to the rise and fall of the melody. The shape of the melody in this piece is indeed as simple as the harmony is complex. In the spirit of the diatonic underpinnings of this chromatic music, let us assign pitch height in imitation of scale degrees (non-modulo 8), as shown in Figure 33 for the second phrase. A diatonic and a lowered sixth degree both receive “6,” for example. (One could also measure pitch height by semitones from the bottom melodic pitch B3, which we have done with comparable results.)

The correlation of melodic contour alone with the data is surprisingly robust: $R^2(1,45) = .66$, $p < .0001$, $R^2_{adj} = .65$. Figure 34 shows the result when contour is combined by multiple regression with the other factors that predict tension: $R^2(3,43) = .67$, $p < .0001$, $R^2_{adj} = .65$; $p(\text{attraction}) = .02$, $\beta = .22$; $p(\text{tension}) = .003$, $\beta = .33$; $p(\text{contour}) < .0001$, $\beta = .57$. Now the correlation is healthy, and all the variables, even attraction, are statistically significant. Yet it should be noted that for the best-fitting solutions for all the diatonic passages analyzed,

attraction appears weaker than tension, as shown by a β value that is consistently less than that for tension. This suggests that the numerical formulation of attraction might be improved.

We have computed the melodic-contour factor for all the other music under consideration and found that it is significant only in the Chopin. Why is this so? The melodic contours in the other pieces move up and down in more intricate patterns than in the Prelude, and their harmonic structures are less complicated. Generally speaking, listeners gravitate toward structures that are more easily processed. When faced with the simple melody and convoluted harmonies of the Chopin, they apparently give the former greater weight than they would under more usual circumstances.

Analyses in Nondiatonic Spaces

Despite its intense chromaticism, the Chopin prelude remains diatonic in the sense that its harmonic progressions refer to diatonic scale degrees, albeit often chromatically inflected degrees. Later composers began to explore progressions that, while still having a tonal orientation, refer to nondiatonic structures, such as the whole-tone scale (all whole steps), the octatonic scale (alternating half and whole steps), and the hexatonic

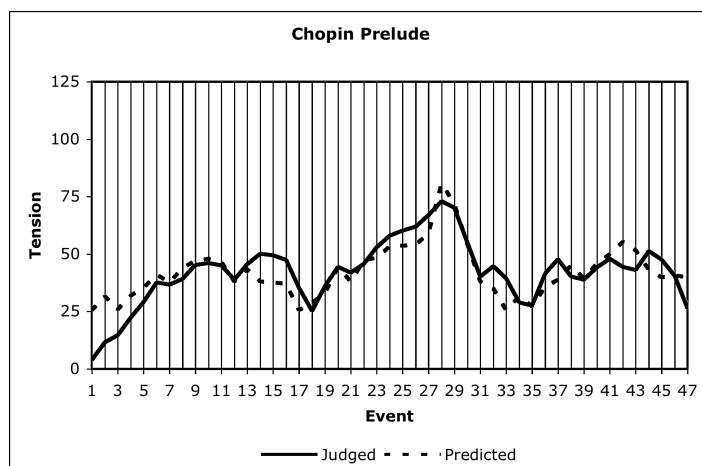


FIGURE 34. Tension graph for the TPS + contour analysis of the Chopin prelude.

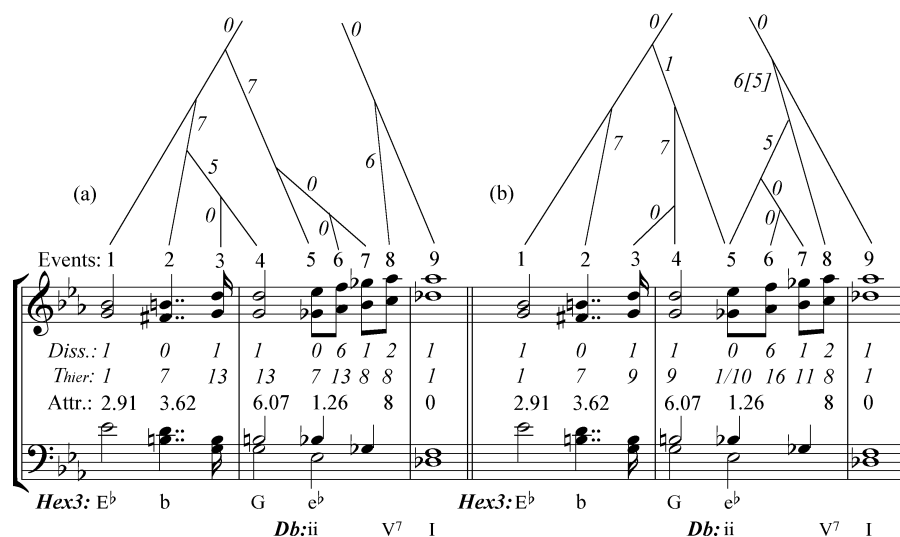


FIGURE 37. Hexatonic-to-diatonic analyses of the chromatic version of the Grail motive: (a) right-branching interpretation; (b) TPS interpretation.

from Event 4 but directly from Event 1, a consequence of the hypermodulation at that point from hexatonic to diatonic space, with $\delta_s(\text{Event 1} \rightarrow \text{Event 5}) = 7$. Figure 38 reveals a poor data fit for this analysis: $R^2(2,6) = .34$, $p = .29$, $R^2_{adj} = .11$; $p(\text{attraction}) = .94$, $\beta = -.0322$; $p(\text{tension}) = .16$, $\beta = .59$. The predictions for Events 3-4 are too high because of the inherited value from Event 2. The prediction for Event 5 is conversely too low; the chromatic progression between the E^b major and minor triads of Events 1 and 5 apparently dilutes the connection between them. The discrepancy at Event 8 does not result not from a local calculation and may be an artifact of the statistical analysis.

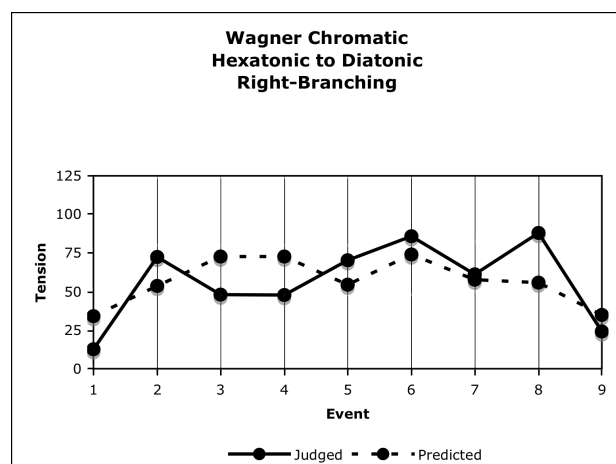


FIGURE 38. Tension graph of the analysis in Figure 37a.

Figure 37b repeats the analysis in Chapter 7 of *TPS*. In keeping with its pivot function, Event 5 takes a double branch, from the right in hexatonic space and from the left in diatonic space. Hence Event 5 has two tension values, 1 and 10; the latter number fits the data better and is taken here. The improved result, shown in Figure 39, gives $R^2(2,6) = .66$, $p = .04$, $R^2_{adj} = .55$; $p(\text{attraction}) = .44$, $\beta = .19$; $p(\text{tension}) = .02$, $\beta = .78$.

In both analyses, attraction probabilities are high, indicating that this factor makes little contribution to the correlation. Moreover, the more successful hexatonic analysis (Figure 37b) bears little resemblance to the best-fit analysis of the diatonic version of the

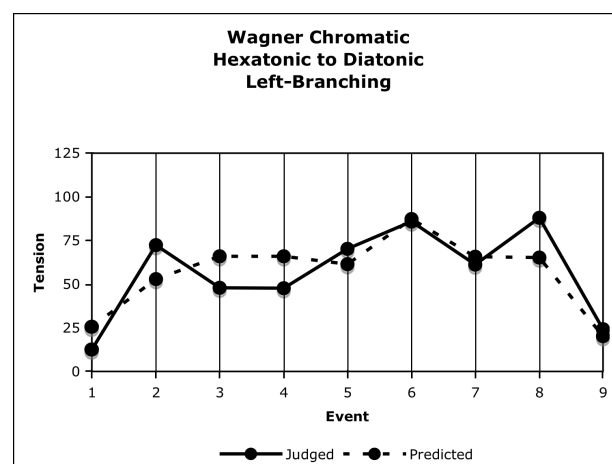


FIGURE 39. Tension graph of the analysis in Figure 37b.

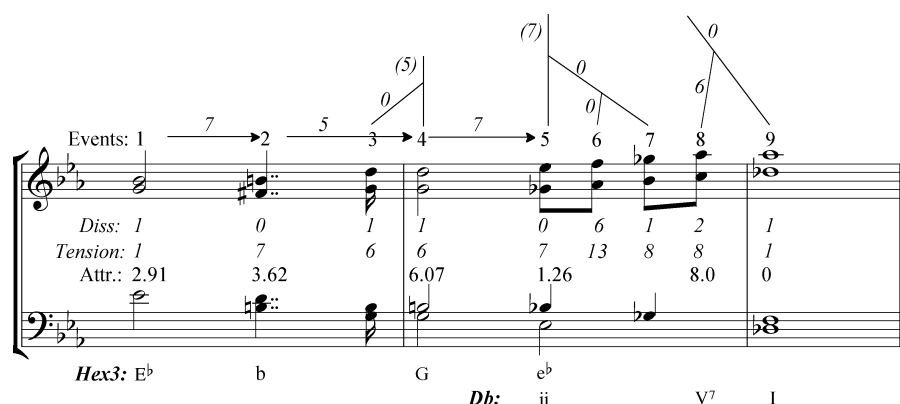


FIGURE 40. Sequential-to-hierarchical, hexatonic-to-diatonic analysis of the chromatic Grail theme.

theme (Figure 22). It is hard to justify such contrasting analyses of the two versions.

The predicted curves diverge from the data curves at the same places in Figures 38 and 39: increase instead of decrease between Events 2 and 3-4, and decrease instead of increase between Events 4 and 5. This pattern suggests a sequential rather than hierarchical approach to these events, for a sequential analysis eliminates inherited values at Events 3-4 and explicitly includes the relatively large distance that exists between Events 4 and 5. A larger rationale behind this step is that, when faced with puzzling chromatic progressions, listeners seem to find a hierarchical construal more difficult, for common schemas are not triggered. Instead the tendency may be to make sense of each event merely in terms of the previous one. Hierarchical hearing becomes robust once the music becomes diatonic at Event 5.

Figure 40 carries out this approach. The arrows for Events 1-5 represent sequential rather than hierarchical distances. (A slight exception is that Event 3 remains a left-branching anticipation of Event 4.) At Event 5, fragments of a prolongational tree emerge, along the lines of the right-branching treatment of Events 5-7 in Figure 22. A left-branching treatment of Event 5, producing an integrated tree for Events 5-9, as in Figure 11, would yield almost the same tension numbers.

This interpretation yields Figure 41. The predicted curve at the beginning of the phrase comes closer to that of the data because the distance from Event 1 to Event 2 is greater than the distance from Event 2 to Event 4. Similarly, the predicted curve from Event 4 to Event 5 improves because of the relatively large distance between them. The overall fit is strong: $R^2(2,6) = .83$, $p = .005$, $R^2_{adj} = .77$; $p(\text{attraction}) = .42$, $\beta = .15$; $p(\text{tension}) = .002$, $\beta = .88$. Attraction does not reach significance, however.

Although Figure 41 is satisfactory, it is possible that a good fit can also be obtained by treating the chromatic Grail theme entirely within diatonic space. Perhaps listeners are so unaccustomed to hexatonic space that they intuitively stick with familiar diatonic space, even if it has to be adjusted to accommodate the chromatic input. Figure 42a gives a hierarchical analysis in which the Roman-numeral designations descend by half step, from Eb to D to Db, in parallel with the flattening of Events 2 and 5. The modulation is sufficiently inexplicit that we decided to suppress the i variable in δ , leaving distance to be measured by chord root distance (j) and noncommon tones (k). As shown in Figure 43, this analysis fails empirically: $R^2(2,6) = .30$, $p = .34$, $R^2_{adj} = .07$; $p(\text{attraction}) = .45$, $\beta = .28$; $p(\text{tension}) = .25$, $\beta = .44$. The outcome bears comparison to the hexatonic-to-diatonic analyses in Figure 37 and the data correlation in Figure 38. Related prolongational variants would yield similar results.

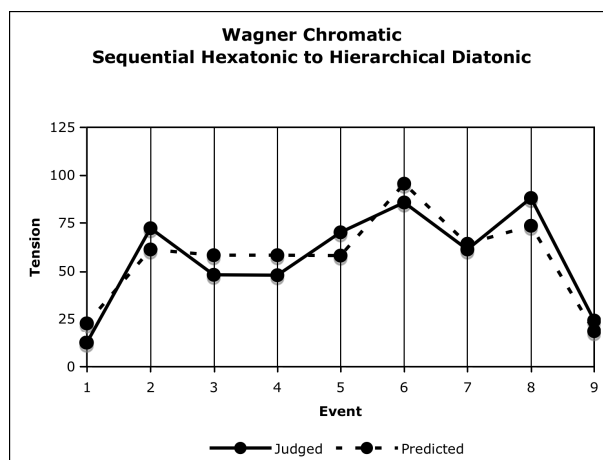


FIGURE 41. Tension graph of the analysis in Figure 40.

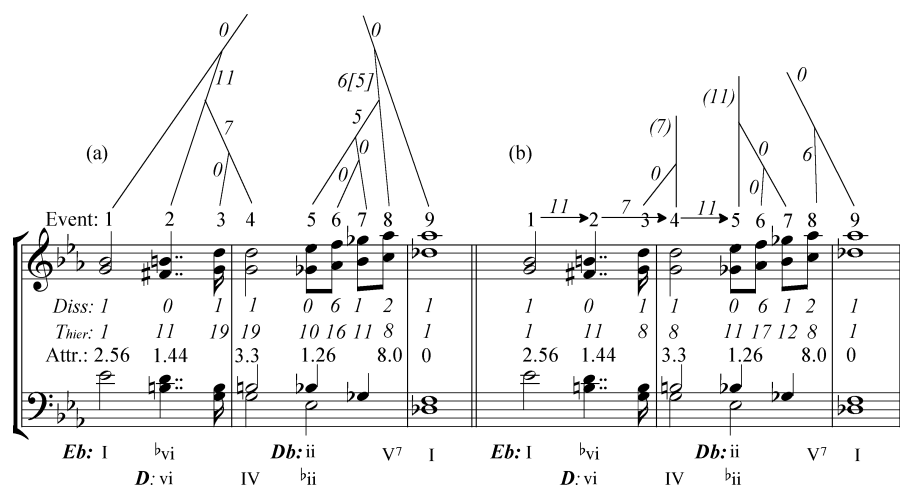


FIGURE 42. Modulating diatonic analyses of the chromatic Grail theme: (a) hierarchical interpretation; (b) sequential to hierarchical interpretation.

Figure 42b takes a sequential-to-hierarchical interpretation akin to Figure 40, except that this time the analysis takes place not partly in hexatonic space but entirely within modulating diatonic space. Again variable i is inactivated for Events 1-5. The distance calculations are from I to bvi in Eb, and vi to IV to bii in D; the latter pivots as ii/Db in diatonic space. The strong correlation, shown in Figure 44, is $R^2(2,6) = .79$, $p = .0025$, $R^2_{adj} = .72$; $p(\text{attraction}) = .04$, $\beta = .28$; $p(\text{tension}) = .0012$.

To summarize this rather complicated discussion of the chromatic version of the Grail theme, the analyses that work best with the data are those that treat the chromatic part of the progression sequentially and the

diatonic part hierarchically. The correlations for the hexatonic-to-diatonic interpretation (Figures 40 and 41) and the shifting-diatonic interpretation (Figures 42b and 44) are too close to decide between them. The force of the latter, however, is weakened by two theoretical factors: the TPS space-finding method picks hexatonic space for Events 1-5, and the shifting-diatonic interpretation relies on a doubtful suppression of the i variable in δ . Even so, the evidence is uncertain whether listeners hear Events 1-5 in hexatonic space or in shifting-diatonic space. (An alternative approach to this passage within a neo-Riemannian framework is presented in Appendix B.)

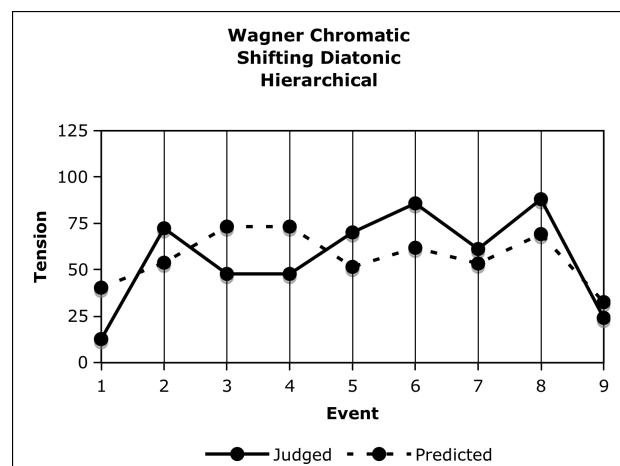


FIGURE 43. Tension graph for Figure 42a.

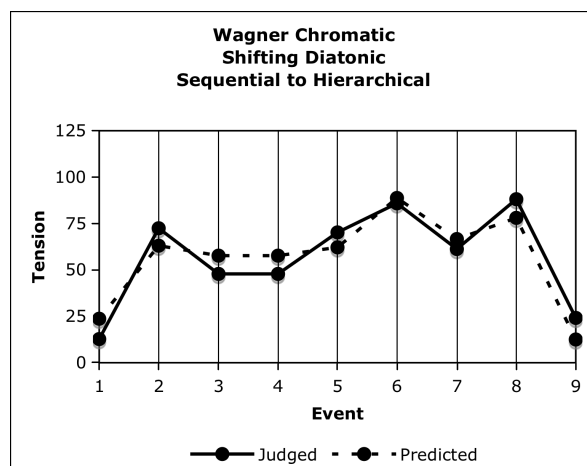
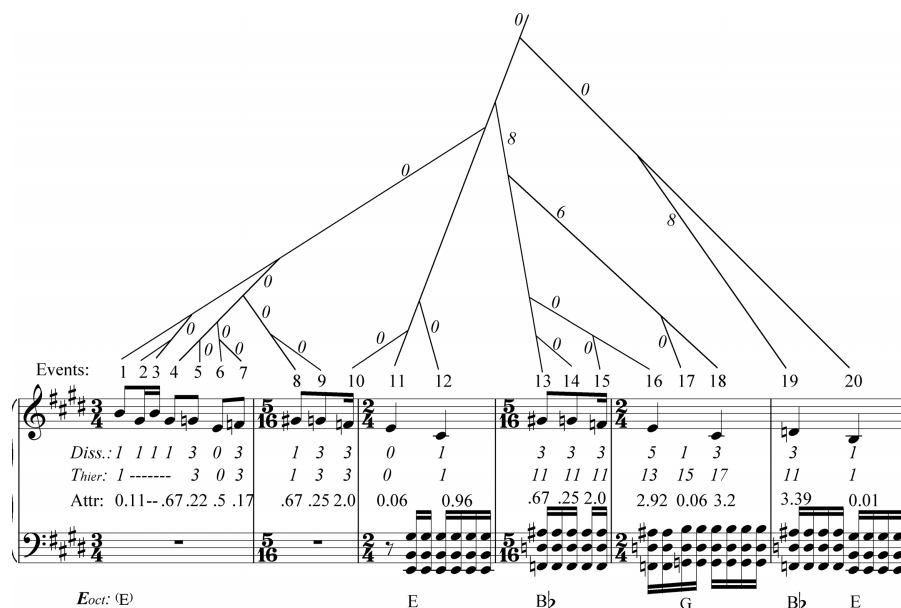


FIGURE 44. Tension graph for Figure 42b.

FIGURE 45. Analysis of the first phrase of Messiaen's *Quartet for the End of Time*, V.

Messiaen Quartet

Figures 45-46 give the opening parallel phrases of the fifth movement ("Louange à l'Eternité de Jésus") from Messiaen's *Quartet for the End of Time*. Events are numbered according to melodic and/or harmonic changes. The melody in the original is played on the cello and the

repeated chords on the piano, but for the experiment both parts were played on the piano. To shorten the experiment slightly, the lengths of Events 12, 18, 20, 32, 38, and 40 were reduced from half to quarter notes. At a more global level than is shown, Event 31 attaches to Event 11. In the original, Events 39-40 continue into a consequent phrase beginning on IV^{#11}/E; however, the

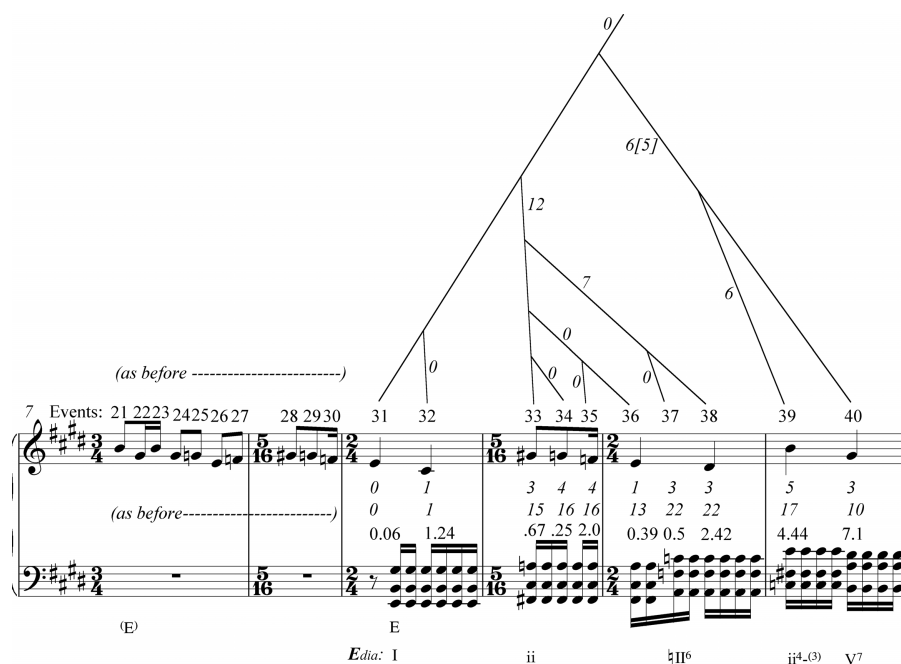


FIGURE 46. Analysis of the second phrase of the Messiaen.

subjects heard the music only up to Event 40. It is therefore assumed in the analysis that Event 40 ends in a half cadence in relation to the tonic E major chord. Events 1-10 are subordinate to Event 11 because they lack explicit harmonic support; similarly for Events 21-30 to Event 31. In both cases, an E major harmony is implied.

All the pitches in Figure 45 and up to the fourth bar of Figure 46 belong to a single octatonic collection, **oct1**, or, equivalently, E_{oct} —that is, an E major tonic over an octatonic scale. (Messiaen, 1944, calls the octatonic scale the second mode of limited transposition.) With the introduction of F# and A at Event 33, the E major of Events 31-32 function retrospectively as a hypermodulatory pivot to E_{dia} (E major tonic in a diatonic context). The interspatial distance rule consequently adds 1 to Events 33-40. The F major chord at Events 37-38 serves as a flattened supertonic in E_{dia} . Within the given harmonic framework, melodic pitches can be nonharmonic in different ways. For example, Events 14-15 are octatonic scale tones outside the E major chord; but at their repetition in the parallel phrase, Events 34-35 are chromatic nonscale tones within an E major diatonic context. The melodic notes in Events 39 and 40 are unresolved appoggiaturas (B implies A, G# implies F#), a characteristic stylistic feature described in Messiaen (1944).

As with the hexatonic interpretation of the chromatic Grail theme, the question arises whether listeners hear the octatonic-to-diatonic interpretation assigned in Figures 45-46 or whether they try to fit the entire passage into a diatonic schema. The tension graph in Figure 47 supports the former interpretation: $R^2(2,37) = .76$, $p < .0001$, $R^2_{adj} = .75$; $p(\text{attraction}) = .0009$, $\beta = .33$;

$p(\text{tension}) < .0001$, $\beta = .67$. As with the Chopin excerpt, the presented tempo was slow (4 seconds per quarter-note beat), so the discrete values labeled “predicted” in the graph were computed from the continuous-tension judgments by averaging the judgments from the onset of each event to the onset of the next event. Again, the continuous-tension task causes a misleading discrepancy for the first few events. Here events 21-25 repeat Events 1-5, however. If the data values for events 21-25 substitute for those of Events 1-5, the excellent result is $R^2(2,37) = .85$, $p < .0001$, $R^2_{adj} = .84$; $p(\text{attraction}) < .0001$, $\beta = .35$; $p(\text{tension}) < .0001$, $\beta = .71$. We note that, as in the case of the diatonic excerpts, the β values for attraction are consistently lower than those for tension; indeed, attraction appears generally weaker relative to tension for the nondiatonic excerpts than for the diatonic excerpts.

In several places the model makes inaccurate predictions. Event 18 is given too high a tension value. The desynchronization of melody and harmony at this point—the G major chord arrives two 16th notes before the beat—probably softens the perception of dissonance when the C# arrives. Perhaps the tension data for Events 26-31 are higher than predicted because of anticipatory tension for the chords to reenter. The relatively high tension perceived at Event 36 may result from the ongoing trajectory of the melody in mid-phrase. As far as we can see, however, any adjustment made in the model to improve these specific instances would produce greater negative consequences elsewhere.

The octatonic-to-diatonic interpretation is almost matched by an entirely diatonic analysis, for which

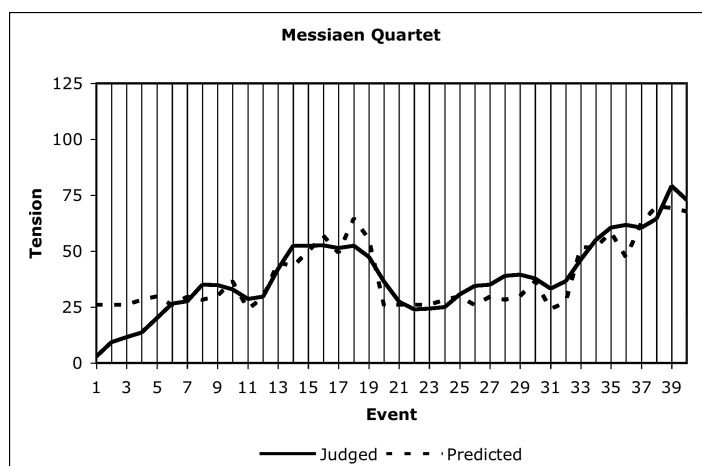


FIGURE 47. Tension graph for the Messiaen.

$R^2(2,37) = .70$, $p < .0001$, $R^2_{adj} = .68$; $p(\text{attraction}) < .0006$, $\beta = .37$; $p(\text{tension}) < .0001$, $\beta = .62$. The tension and attraction numbers in the two interpretations are in fact similar, so there is no need to display the graph for the diatonic version. Because of this convergence, the evidence for octatonic listening is not decisive. Yet the evidence for octatonic listening in the Messiaen is slightly stronger than for hexatonic listening in the Wagner. We surmise two reasons for this. First, the octatonic collection appears more often, and in more kinds of music (such as jazz), than does the hexatonic collection. Greater exposure facilitates the growth and use of the octatonic basic space. Second, the octatonic passage is much longer than the hexatonic one, and it is articulated melodically as well as harmonically. Listeners have time to get used to an octatonic framework. Additional support for octatonic hearing was found in a probe-tone study of Stravinsky's Petroushka chord (Krumhansl & Schmuckler, 1986).

Discussion

We began with the claim that four components are required for a quantitative theory of tonal tension: prolongational structure, the pitch-space model, an approximation of surface dissonance, and the attraction model. Now, having tested the theory with empirical judgments of musical tension with a variety of excerpts, we reflect on the implications of the results for both the theory and the experience of tension. The degree to which the theoretical predictions were confirmed by the empirical data provides strong support for the theory. The analyses presented here also illuminate different aspects of the theoretical components. We consider each in turn.

The first component of the theory, prolongational structure, asserts that listeners hear music hierarchically. That is, listeners understand each sounded event not just in relation to immediately preceding and following events but also in terms of nonadjacent dependencies. The hierarchy is represented by a tree structure that specifies the embedding of events. Using distances from the pitch-space model, tonal tension is calculated down the branches, so that events inherit distances from events superordinate to them. The degree to which the data supports this claim has significant implications for music cognition. They support the existence of complex cognitive representations in which events are encoded and remembered with respect to events that occur at a distance, either in the past or anticipated to occur in the future.

Given the importance of this claim for a cognitive theory of music, we emphasize that the sequential

treatments include all the complexities of the other three components of the *TPS* theory—pitch space distances, surface dissonance, and attraction. We have demonstrated the specific ways in which the sequential model fails to account for the diatonic version of the Wagner *Grail* theme. We have considered purely sequential and mixed sequential and hierarchical models for other excerpts as well, although the details of those analyses are not reported here. The results of these tests consistently point to the greater efficacy of the hierarchical models.

The success of the hierarchical tension model in fitting the tension ratings does not mean that sequential listening is discounted. Sequential listening is assumed for the attraction component, in which attractions are computed in a strictly sequential manner. To the extent that attraction contributes to the overall tension, the model reflects a balance of hierarchical and sequential listening.

Earlier attempts to test the empirical validity of theories of pitch reduction, whether in Schenker's (1935) version or that of *GTTM*, have encountered difficulties in accessing the relevant intuitions (for example, Dikken, 1994; Serafine, 1987). Our research offers a fresh way to investigate the psychological efficacy of pitch-reduction theories, not through systematic modifications of the musical surface (as in Dikken) but indirectly through the measurement of tonal tension. Listeners' unpremeditated awareness is not of hierarchical structure per se but of the patterns of tension and relaxation that arise from it. Theory speaks in terms of, say, the composing out of a tonic prolongation, but the immediate experience is one of a rise and fall in tension. As a result, the empirical validity of different prolongational analyses can be explored through the tension patterns they represent.

Given the complexity of the tension model, and especially of its prolongational component, it may be asked if a more parsimonious model could achieve similar empirical results. There have been attempts, for example, to model tonal cognition by neural networks (Bharucha, 1987; Leman, 1995; Tillmann, Bigand, & Bharucha, 2000). Such models have the dual advantages of being purely computational, arriving at representations directly from the music without external theoretical claims, and of offering plausible hypotheses about the learnability of the resulting representations. While these studies have partly succeeded in recovering circle-of-fifths relations and major-minor tone profiles, they do not begin to address the fine distinctions in tension and relaxation that are elicited in the course of listening to music of any length or intricacy.

In particular, it is not clear how neural network models could describe the hierarchical relations embodied in the prolongational component of *TPS*. One possibility would be for the neural network models to analyze music on different time scales; local relations would be expected to appear on shorter time scales, whereas global, hierarchical relations might emerge over longer time scales. It would be of considerable interest if such models were able to make predictions and produce correlations with empirical data comparable to those of the tonal tension model. These models could then be evaluated on competing claims of complexity, hierarchy, learnability, and domain-specificity.

The empirical research that comes closest to the present study is Bigand and Parncutt (1999). Their second experiment employs a chord sequence close to the Chopin sequence in Figures 29-30. Using a methodology that resembles our stop-tension task, they find that, contrary to the present study, listeners hear almost entirely sequentially rather than hierarchically. To a degree, this outcome can be explained by the absence in their model of the components of surface dissonance and attraction. In the second phrase, for example, their data shows greater relaxation at the arrival of $I/A\flat$ (Event 31 in Figure 30) than at the ensuing V^7/E (Event 32) that leads back to the global tonic. Their model would have predicted greater tension for V^7/E if it had included surface dissonance and, especially, an attraction component, since the expectation for I/E is strong at that point. Other details in their data, however, are puzzling, such as the increase in tension at the arrivals of I/E at Events 17 and 33. It is possible that their methodology encouraged moment-to-moment listening. The instructions given to listeners were to judge an event as tense if it evoked nonclosure, that is, the "feeling that there must be a continuation of the sequence" (Bigand & Parncutt, p. 242). However, closure/non-closure does not necessarily correspond to stability/nonstability. Closure typically occurs at the end of a phrase, but a stable event may initiate a phrase and hence the urge for continuation. Such is the case with Events 17 and 33. Moreover, immediate closure can be quite powerful if a local tonic is approached by left branching, even if that tonic is relatively unstable in the larger picture. Such is the case with Event 31 ($I/A\flat$). It appears that their data partly reflect intuitions other than those of the tension and relaxation of stable and unstable events. In any case, our data contradict their claim that "musical events are perceived through the frame of a short window sliding along a sequence, so that events perceived at a given time are negligibly influenced by events outside the window" (Bigand & Parncutt, p. 252).

The support found in our research for prolongational structure suggests a deep parallel between music and language. Nonadjacent dependencies are held to be a special aspect of language structure (Chomsky, 1957; Hauser, Chomsky, & Fitch, 2002). In the sentence "The man who left forgot his black umbrella," the subordinate clause "who left" requires that a link be made between the nonadjacent elements "man" and "forgot." Considerable interest has focused recently on the issue of whether processing such hierarchical structures is unique to humans (Fitch & Hauser, 2004; Hauser, Newport, & Aslin, 2001; but see Perruchet & Rey, 2005). Other research has explored the conditions under which humans, especially infants, have the ability to learn nonadjacent dependencies in artificial languages (Gomez, 2002; Gomez & Maye, 2005; Newport & Aslin, 2004). Additionally, researchers have found that neural network models can learn such structures (Christiansen & Chater, 1999) and that unique neural resources may be devoted to processing hierarchical structures in language (Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006). The present results supporting hierarchical processing suggests that although the ability may prove unique to humans, it is not limited to language.

An important difference between musical and linguistic hierarchies should be noted. The prolongational tree, unlike a syntactic or phonological tree, expresses tensing and relaxing relations. As discussed, if Event 1 tenses into Event 2, the analysis assigns a right-branching structure, indicating that Event 2 is an elaboration of Event 1. If, however, Event 1 relaxes into Event 2, the result is a left-branching structure, indicating that Event 1 is an elaboration of Event 2. This distinction codifies the central role that the ebb and flow of tension plays in the experience of music. Right and left branching also appears in linguistic analyses, but only to signify that the subordinate element is before or after the dominating element (the "head" in X-bar phrase-structure grammar; see Jackendoff, 1977). In the sentence above, the subordinate clause "who left" right-branches from the preceding "man"; "black" left-branches to "umbrella." These distinctions do not carry the implication of increasing or decreasing tension.

Another difference arises from the comparison between the musical analyses presented here and the language studies cited above that involve artificial languages. In the latter, dependencies are arbitrarily established in the learning corpus. In music, dependencies build on previously established hierarchical relationships between tones and chords as described in abstract tonal hierarchies. Certain pitches and chords are more

stable in a tonal system and serve as reference points (see *CFMP*). For example, the sounding of a V chord engenders the expectation that a I chord will follow. In the musical analyses, the inherited tension of an event is computed from the more stable events superordinate in the prolongational tree. The experience of tension, as theorized in *TPS* and reflected in listeners' judgments, draws on prior knowledge of the abstract hierarchies, not just dependencies established within the particular musical excerpt. In this sense, the musical case is more like natural language processing involving prior syntactic and semantic knowledge. For example, hearing "if" in a sentence induces the expectation for a later "then" in the sentence. It is an open question how well listeners are able to process long-distance dependencies in music established without the support of such prior stylistic knowledge.

Although hierarchical analyses were supported in most of the cases considered, two qualifications are suggested by the results. The first is that listeners have a tendency to hear retrospectively rather than prospectively. That is, unless they have strong schematic expectations, they are likely to interpret events with respect to events already heard. This favors right- over left-branching prolongational relations. This preference was apparent, for example, in the diatonic version of Wagner's Grail theme. The second qualification is that instances of sequential hearing can arise. Such a case is the chromatic version of the Grail theme, which begins with the relatively unfamiliar hexatonic scale collection and unusual chromatic progressions. These conditions suggest that listeners tend to fall back on a mode of sequential processing when the musical surface fails to engage familiar tonal patterns.

A final point concerning prolongational structure is that most of the revisions in our analyses occurred within in this component of the theory. For each of the musical excerpts, a variety of prolongational variants were considered, only some of which were reported here. These reevaluations reflect the fact that although prolongational analyses are constrained by general principles and specific formalisms, there is small leeway in their application to specific musical passages. This is because, unlike the other components, which are strictly algorithmic, the prolongational component involves gradient preference rules, as described in *GTTM* and *TPS*. Individual rules describe conditions that weigh in favor of one or another analysis, but the different rules often conflict, so that in most cases a single condition is not decisive. The empirical data can play a significant role in determining an optimal solution, and may offer insight into the relative rankings of

preference rules, potentially making possible an algorithmic formulation of the prolongational component.

We were able to apply the second component of the theory, the pitch-space model, without need for revision. This component describes the calculation of the distance between any chord in one region (or key) and any other chord in that region or any other region. Is it surprising that this quantification was strongly supported by the empirical tests? In one sense, no, because the pitch-space model developed in *TPS* draws heavily on previous perceptual experiments that measure distances of pitches, chords, and regions, as summarized in *CFMP*. Yet there are reasons to think that difficulties might have been encountered when these distances were incorporated into the *TPS* theory. One reason is that most of the empirical data on tone, chord, and key distances came from experiments using short, schematic stimulus materials such as scales and chord cadences. In addition, the distance judgments were obtained for pitches and chords presented in immediate succession. It was not obvious in advance that the same distances would apply to the considerably more complex musical materials discussed here. Nor was it clear that the same distances would apply in the context of prolongational structures, which often specify long-range dependencies between nonadjacent events. In these ways, then, the resiliency of the distance measures as incorporated in the *TPS* theory is notable.

Another important point about the pitch-space component is that the tests supported extensions to nondiatonic scale systems. These extensions adapt the three subcomponents of pitch space—pitch, chord, and key distances—to other tonal frameworks. The two systems considered here were the hexatonic scale (for the beginning of the chromatic version of the Grail theme) and the octatonic scale (for most of the passage from Messiaen's *Quartet for the End of Time*). The relatively successful match to the empirical data for these passages shows that experimental and theoretical measures of distances that were developed initially for diatonic music may apply as well to chromatic tonal systems that have received less empirical study, thus suggesting a degree of generality for these measures.

Finally, although no comparably quantified theoretical model is currently available for comparison, we examined how the pitch-space distances of *TPS* might function within a neo-Riemannian framework (see Appendix B). Despite other differences between the *TPS* and neo-Riemannian theories, they derive essentially the same geometric descriptions of octatonic and hexatonic spaces. One problem, then, was to devise a suitable measure of chord distance for the

neo-Riemannian analysis. Central to the theory are three transformations: L (leading-tone), P (parallel), and R (relative). The distance calculation for chords was taken to be the minimal number of transformations needed to move from one chord to the next. The other modification was to reduce the basic pitch space to two levels (chromatic and triadic) because the neo-Riemannian approach does not incorporate scales or a hierarchy internal to the triad (root, fifth, and third). With these modifications, the neo-Riemannian predictions matched the empirical tension ratings for the chromatic Wagner theme just as well did as the best-fitting *TPS* (largely sequential) analysis. Neo-Riemannian predictions for the diatonic version of the theme did not fare very well, however. The diatonic version requires hierarchical treatment beyond the scope of neo-Riemannian theory.

We turn next to the third component of the theory, the measure of surface dissonance. Dissonance is both a psychoacoustic phenomenon that can be continuously quantified and an abstraction suited to discrete categorization. The *TPS* approach combines the two by using discrete categories based on the underlying psychoacoustics. It determines if a melodic pitch is on the third or fifth scale degree, if the chord is in inversion, or if there is a nonharmonic tone. Quantitative values are assigned and summed together to yield an overall measure of surface dissonance. This approach may not take full advantage of psychoacoustic measures (which would ideally take into account register, timbre, dynamics, and other factors), but it reflects the intuition that dissonance is experienced to some extent in different categorical types.

Moreover, the proposed measure of surface dissonance is largely supported by the empirical tension data. In all but one case, it was not necessary to include other surface features that might contribute to tension, such as note density, dynamics, and pitch range (as found, for example, in the tension study of Krumhansl, 1996). The exception was the Chopin Prelude. Only when we added melodic contour was a satisfactory analysis identified. Two characteristics of the Chopin might have contributed to this result. First, its melodic contour is clear and simple. Second, the contour is set against a progression of exceedingly complex, chromatic harmonies. These conditions apparently direct the listeners' attention to the melodic contour because the melody is much easier to process than is the harmonic progression. In general, however, the perceptual data and the theoretical analyses have successfully focused on the sense of tension created by melodic and harmonic motion specific to music and independent of

contour, as opposed to factors such as dynamics and range that are shared with other domains such as speech, gesture, and dance.

The fourth and last component of the theory is attraction. This, like the second and third components, is calculated algorithmically. It represents the intuition of the pull or urge of events (pitches and chords) toward other events, especially those that are more stable in the tonal context. This phenomenon has a strong empirical basis in similarity ratings and can account for patterns of errors in memory, as summarized in *CFMP*, Chapter 5, and as described by the three principles of tonal stability in *CFMP*, Chapter 6. (Also see Bharucha, 1984, 1996; Bharucha & Krumhansl, 1983.) It has also been extensively theorized (Larson, 1994, 2004; Margulis, 2005; *TPS*) and appears to be closely related to melodic expectancy (Meyer, 1956; Narmour, 1990), another topic that has been studied in numerous experiments.

Despite this wealth of research, issues remain about how best to quantify attraction. One issue is how strongly the distance between pitches affects the degree to which two tones are attracted to one another. Another issue is the best measure of stability. In the context of *TPS*, it is a question of how many levels should be included in the basic space. A third issue concerns the influence on attraction values by relations at multiple levels of the tonal hierarchy (or alphabets; see Deutsch, 1980). For example, if attractions included calculations at the triadic level of the *TPS* basic space, this would increase the attraction between triad members. Another alternative, considered in *TPS*, is whether attractions might also be computed from event to event not only at the surface but also at the immediately underlying levels of prolongational reduction. For the present, we have retained the simple algorithm proposed in *TPS*, which has proved serviceable for the applications presented here.

A final observation about the quantification of attraction is of interest. As mentioned, the approach of *TPS* was extended to neo-Riemannian theory and tested with the chromatic version of the Grail theme. This extension reduced the basic space to only two levels. Despite this, the resulting attraction values made a contribution to the fit of the data, although its influence was relatively weak. In general, our tests showed that, as a predictor, attraction was consistently weighted less than tension. For several reasons, then, we have less confidence in the quantification of attraction than in the other two algorithmic components of the theory.

Further progress in improving and testing the four components of the tension model depends on its

implementation as a computer program that performs automatic analysis and permits interactive options. Doing the calculations by hand is laborious and liable to clerical error. In addition to easing calculations, implementation of the model would facilitate the investigation of analytical options and make possible the study of many more musical passages, including much longer ones. It would be of great interest to have a theoretical and empirical analysis of tonal tension over long time spans. Needless to say, the development of an automatic, interactive analysis system is in itself a daunting task.

Another extension is to musical styles beyond those considered here. The excerpts in this article represent musical syntax from eighteenth-century diatonic tonality to twentieth-century nondiatonic tonality. *TPS* traces this evolution through changes in the structure of pitch space and operations upon it. A similar treatment could be tested with older tonality, going back to music of the Renaissance and Middle Ages. The tension model, adapted suitably, can help explain fundamental features in a given tonal style and, ultimately, in the evolution of

tonality in terms of fundamental perceptual and cognitive principles.

Author Note

This research was supported in part by a fellowship grant from the J. Simon Guggenheim Foundation to the second author. The authors acknowledge the assistance of Erin Hannon in conducting the experimental sessions for the experiment with the Bach and Wagner excerpts. Helpful comments by David Temperley, William Thompson, Dmitri Tymoczko, and an anonymous referee improved the contents of this paper.

Correspondence concerning this article should be addressed to Fred Lerdahl, Department of Music, 621 Dodge Hall, Mail Code 1813, Columbia University, New York, NY 10027, or to Carol L. Krumhansl, Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853. E-MAIL: awl1@columbia.edu (Lerdahl) and clk4@cornell.edu (Krumhansl).

References

-
- ALDWELL, E., & SCHACHTER, C. (1979). *Harmony and voice leading*, Vol. 2. New York: Harcourt Brace Jovanovich.
- BERRY, W. (1976). *Structural functions in music*. Englewood Cliffs, NJ: Prentice-Hall. (Reprinted New York: Dover, 1987.)
- BHARUCHA, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, 16, 485-518.
- BHARUCHA, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, 5, 1-30.
- BHARUCHA, J. J. (1996). Melodic anchoring. *Music Perception*, 13, 383-400.
- BHARUCHA, J. J., & KRUMHANSL, C. L. (1983). The hierarchical representation of harmonic structure in music. *Cognition*, 13, 63-102.
- BIGAND, E., & PARNCUTT, R. (1999). Perception of musical tension in long chord sequences. *Psychological Research*, 62, 237-254.
- BIGAND, E., PARNCUTT, R., & LERDAHL, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, 58, 124-141.
- CHOMSKY, N. (1957). *Syntactic structures*. The Hague: Mouton.
- CHRISTIANSEN, M. H., & CHATER, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, 23, 157-205.
- CLAMPITT, D. (1998). Alternative interpretations of some measures from *Parsifal*. *Journal of Music Theory*, 42, 321-341.
- COHN, R. (1996). Maximally smooth cycles, hexatonic systems, and the analysis of late-Romantic triadic progressions. *Music Analysis*, 15, 9-40.
- COHN, R. (1998). Introduction to neo-Riemannian theory: a survey and a historical perspective. *Journal of Music Theory*, 42, 167-180.
- CUDDY, L. L., & SMITH, N. A. (2000). Perception of tonal pitch space and tonal tension. In D. Greer (Ed.), *Musicology and sister disciplines: Past, present, future* (pp. 47-59). Oxford: Oxford University Press.
- DEUTSCH, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics*, 28, 381-389.
- DIBBEN, N. (1994). The cognitive reality of hierarchic structure in tonal and atonal music. *Music Perception*, 12, 1-26.
- FITCH, W. T., & HAUSER, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303, 377-380.
- FORTE, A. (1973). *The structure of atonal music*. New Haven: Yale University Press.
- FRIEDERICI, A. D., BAHLMANN, J., HEIM, S., SCHUBOTZ, R. I., & ANWANDER, A. (2006). The brain differentiates human

- and nonhuman grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences*, 103, 2458-2463.
- GOMEZ, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13, 431-436.
- GOMEZ, R. L., & MAYE, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy*, 7, 183-206.
- HAUSER, M. C., CHOMSKY, N., & FITCH, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569-1579.
- HAUSER, M. D., NEWPORT, E. L., & ASLIN, R. N. (2001). Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton-top tamarins. *Cognition*, 78, B53-B64.
- HINDEMITH, P. (1937/1942). *The craft of musical composition*. Vol. 1. New York: Belwin-Mills.
- HOOKE, J. (2002). Uniform triadic transformations. *Journal of Music Theory*, 46, 57-126.
- HUTCHINSON, W., & KNOPOFF, L. (1978). The acoustical component of Western consonance. *Interface*, 7, 1-29.
- JACKENDOFF, R. (1977). *X-bar syntax: A study of phrase structure*. Cambridge, MA: MIT Press.
- KLUMPENHOUWER, H. (2005). Review of *Tonal Pitch Space*. *Journal of the American Musicological Society*, 58, 488-496.
- KRUMHANS, C.L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11, 346-384.
- KRUMHANS, C.L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
- KRUMHANS, C.L. (1996). A perceptual analysis of Mozart's Piano Sonata, K. 282: Segmentation, tension and musical ideas. *Music Perception*, 13, 401-432.
- KRUMHANS, C.L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Psychology*, 51, 36-52.
- KRUMHANS, C.L. (1998). Perceived triad distance: Evidence supporting the psychological reality of neo-Riemannian transformations. *Journal of Music Theory*, 42, 265-281.
- KRUMHANS, C.L., & KESSLER, E.J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial map of musical keys. *Psychological Review*, 89, 334-368.
- KRUMHANS, C.L. & SCHMUCKLER, M.A. (1986). The *Petroushka* chord: A perceptual investigation. *Music Perception*, 4, 153-184.
- KURTH, E. (1920). *Romantische Harmonik und ihre Krise in Wagners "Tristan"* [Romantic harmony and its crisis in Wagner's *Tristan*]. Bern: Haupt. (Reprinted Hildesheim: Olms, 1975).
- LARSON, S. (1994). Musical forces, step collections, tonal pitch space, and melodic expectation. *Proceedings of the Third International Conference on Music Perception and Cognition*, 227-229.
- LARSON, S. (2004). Musical forces and melodic expectations: Comparing computer models and experimental results. *Music Perception*, 21, 457-98.
- LARSON, S. & VANHANDEL, L. (2005). Measuring musical forces. *Music Perception*, 23, 119-136.
- LEMAN, M. (1995). *Music and Schema Theory: Cognitive Foundations of Systematic Musicology*. Berlin: Springer.
- LERDAHL, F. (1988). Tonal pitch space. *Music Perception*, 5, 315-350.
- LERDAHL, F. (1996). Calculating tonal tension. *Music Perception*, 13, 319-363.
- LERDAHL, F. (2001). *Tonal pitch space*. New York: Oxford University Press.
- LERDAHL, F., & JACKENDOFF, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- LERDAHL, F., & KRUMHANS, C. L. (2004). La teoría de la tensión tonal y sus consecuencias para la investigación musical. In J. Martín Galán & C. Villar-Taboada (Eds.), *Los últimos diez años en la investigación musical* [Musical investigation in the last ten years]. Valladolid: Servicio de Publicaciones de la Universidad de Valladolid.
- LEWIN, D. (1984). Amfortas' prayer to Titirel and the role of D in *Parsifal*: The tonal spaces of the drama and the enharmonic C^b/B. *19th-Century Music*, 7, 336-349.
- LEWIN, D. (1987). *Generalized musical intervals and transformations*. New Haven: Yale University Press.
- MARGULIS, E. H. (2005). A model of melodic expectation. *Music Perception*, 22, 663-714.
- MESSIAEN, O. (1944). *Technique de mon langage musical* [Technique of my musical language]. Paris: Alphonse Leduc.
- MEYER, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.
- NARMOUR, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: University of Chicago Press.
- NEWPORT, E. L., & ASLIN, R. N. (2004). Learning at a distance I. Statistical learning of nonadjacent dependencies. *Cognitive Psychology*, 48, 127-162.
- PALMER, C. (1996). Anatomy of a performance: Sources of musical expression. *Music Perception*, 13, 433-453.
- PARNCUTT, R. (1989). *Harmony: A psychoacoustical approach*. Berlin: Springer-Verlag.
- PERRUCHET, P., & REY, A. (2005). Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychonomic Bulletin & Review*, 12, 307-313.
- POVEL, D.-J. (1996). Exploring the elementary harmonic forces in the tonal system. *Psychological Research*, 58, 274-283.
- RIEMANN, H. (1893). *Vereinfachte Harmonielehre, oder die Lehre von den tonalen Funktionen der Akkorde* [Harmony simplified, or instruction in the tonal functions of chords]. (Tr. H. Bewerunge, London: Augener, 1896).
- ROSEN, C. (1972). *The classical style*. New York: Norton.

- ROTHFARB, L. (2002). Energetics. In T. Christensen, ed., *The Cambridge history of music theory*. Cambridge: Cambridge University Press.
- SAMPLASKI, A. (2002). Review of *Tonal Pitch Space*. *Indiana Theory Review*, 23, 185-201.
- SCHENKER, H. (1935/1979). *Free composition*. Tr. E. Oster. New York: Longman.
- SCHOENBERG, A. (1911/1978). *Theory of harmony*. Tr. R. Carter. Berkeley: University of California Press.
- SCHOENBERG, A. (1954/1969). *Structural functions of harmony*, rev. ed. New York: St. Martin's Press.
- SCHOENBERG, A. (1975). *Style and idea*. New York: St. Martins Press.
- SERAFINE, M. L. (1987). *Music as cognition: The development of thought in sound*. New York: Columbia University Press.
- SMITH, N. A., & CUDDY, L. L. (2003). Perceptions of musical dimensions in Beethoven's Waldstein sonata: An application of tonal pitch space theory. *Musicae Scientiae*, 7, 7-34.
- TEMPERLEY, D. (2001). *The cognition of basic musical structures*. Cambridge, MA: MIT Press.
- TILLMANN, B., BIGAND, E., & BHARUCHA, J.J. (2000). Implicit Learning of Tonality: A Self-Organizing Approach. *Psychological Review*, 107, 885-913.
- TOIVIAINEN, P., & KRUMHANSL, C. L. (2003). Measuring and modeling real-time responses to music: Tonality induction. *Perception*, 32, 741-766.
- WEBER, G. (1817-21). *Versuch einer geordneten Theorie der Tonsetzkunst* [Theory of musical composition]. Mainz: B. Schotts Söhne.
- ZUCKERKANDL, V. (1956). *Sound and symbol*. Tr. W. R. Trask. Bollingen Series 44. Princeton: Princeton University Press.

Appendix A: Experimental Procedures

Experimental Method for the Wagner and Bach Excerpts

Participants

Fourteen musically experienced listeners from the Cornell University community participated in the experiment. On average, they had a total of 15.6 years instruction on musical instruments and voice. They had taken an average of 3.0 years of music theory courses and 1.2 years of other music courses at the university level. On a seven-point scale (7 = very familiar), they rated their familiarity with the music of J. S. Bach at 5.5, familiarity with "Christus, der ist mein Leben" at 2.5, familiarity with the music of Wagner at 3.5, and familiarity with *Parsifal* at 2.6. They were paid for participating in the experiment.

Equipment and Stimulus Materials

The experiment was run on a G3 Apple computer with MAX software. The excerpts used were those notated in the text. The sounds were synthesized piano sounds created by Unity, and all tones were played at equal loudness. The music was played over headphones at a comfortable loudness level.

Tasks

Each listener performed two tasks with each of the Bach and Wagner selections. In the stop-tension task, the first event was sounded and listeners judged its

tension. Then, the first two events were sounded, and the listeners judged the tension of the second event. Then, the first three events were sounded and the listeners judged the tension of the third event. This procedure was repeated until all events in the excerpt were sounded. In the continuous-tension task, the listener used a computer mouse to manipulate the position of a slider displayed on the screen to indicate the amount of tension they experienced. The position of the slider was recorded every 100 msec. The stop-tension task was performed before the continuous-tension task.

Procedure

Participants practiced the two tasks with a short excerpt (Bach's chorale "Ermuntre dich, mein schwacher Geist") before beginning the experiment. All listeners judged the Bach excerpt first, initially doing the stop-tension task (hearing the excerpt one time through) and then the continuous-tension task (hearing the excerpt four times through). This means that when they performed the stop-tension task, they had not heard the music beyond the chord they were judging. They then did the two tasks in the same order with the Wagner excerpts, with half the participants judging the diatonic version first and half the participants judging the chromatic version first. At the end of the experiment, participants completed a short questionnaire about their musical experience. The experiment lasted approximately one hour in duration.

Experimental Method for the Chopin and Messiaen Excerpts

Participants

Fifteen musically experienced listeners from the Cornell University community participated in the experiment. On average, they had a total of 19.6 years instruction on musical instruments and voice. They had taken an average of 1.6 years of music theory courses and 2.0 years of other music courses at the university level. On a seven-point scale (7 = very familiar), they rated their familiarity with the music of Chopin at 3.9, familiarity with the E major Prelude at 3.0, familiarity with the music of Messiaen at 2.0, and familiarity with the *Quartet for the End of Time* at 1.8. They were paid for participating in the experiment.

Equipment and Stimulus Materials

The same equipment was used as in the experiment with the Bach and Wagner excerpts. The excerpts used were those notated in the text. All tones were played at an equal loudness.

Tasks

Each listener performed two tasks with each of the Chopin and Messiaen selections. In the continuous-tension task, the listener used a computer mouse to manipulate the position of a slider displayed on the screen to indicate the amount of tension they experienced. The position of the slider was recorded every 100 ms. The participants also performed a continuous-probe task (Toiviainen & Krumhansl, 2003) with these excerpts; those data are not considered here. Because of the inclusion of the continuous-probe task, these two excerpts were played metronomically and at slow tempos.

Procedure

Participants practiced the two tasks with a short excerpt (the Grail theme from Wagner's *Parsifal*) before beginning the experiment. Half the listeners judged the Chopin excerpt first, and half of the listeners judged the Messiaen excerpt first. At the end of the experiment, participants completed a short questionnaire about their musical experience. The experiment lasted approximately one hour in duration.

Appendix B: Neo-Riemannian Theory and Tonal Tension

The chromatic version of the Grail theme is prominent in the neo-Riemannian music-theoretic literature (Clampitt, 1998; Cohn, 1996; Lewin, 1984). How well does a neo-Riemannian analysis predict tonal tension? The question cannot be answered directly. Neo-Riemannian theory (hereafter N-R theory) stems less from traditional tonal theory than from a combination of Riemannian theory (Riemann, 1893), atonal pitch-class set theory (Forte, 1973), and transformational theory (Lewin, 1987). (For an overview of N-R theory, see Cohn, 1998.) The approach does not explicitly address tonal tension or attraction, nor does it incorporate basic elements needed for such a concern. Specifically, it does not include the concepts of region, modulation, chord root, scale, or nonharmonic tone. Even the progression of a triad into a seventh chord (or the reverse) is problematic. Furthermore, N-R theory is designed more for chromatic than diatonic tonal music. Its concern is with

formal operations that transform triads into other triads and the geometries that result from such operations.

Despite these limitations, there are nontrivial points of contact between N-R theory, *TPS*, and *CFMP*. N-R theory and the *TPS* theory generate more or less equivalent geometrical projections of octatonic and hexatonic spaces. Krumhansl (1998) demonstrated that N-R operations generate chord distances comparable to the empirical results reported in Krumhansl and Kessler (1982) and *CFMP*. It will be informative to adapt N-R theory to make tension predictions and to compare the results with those of the *TPS* model.

At the heart of the N-R approach are the L (leading-tone), P (parallel), and R (relative) transformations. All three convert major into minor triads, or vice versa. L lowers the root of a major triad down a half step or the fifth of a minor triad up a half step, as in Figure 48a. (Although the theory does not include scale degrees, it

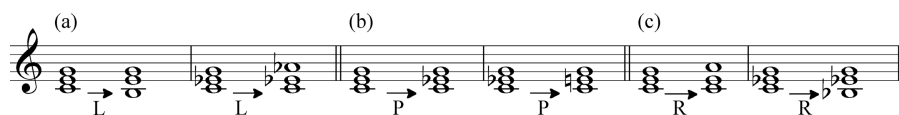


FIGURE 48. L, P, and R operations in N-R theory.

$$\begin{array}{cccccccccccc} & & \textcolor{blue}{3} & & & & \textcolor{red}{7} & & & & \textcolor{green}{10} & & & & \textcolor{blue}{\underline{2}} & & & & \textcolor{red}{6} & & & & \textcolor{green}{\underline{11}} \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ \delta_{nr}(E_b \rightarrow b) = (1 + 2 + 0) + 3 = 6 \end{array}$$

is useful to refer to them when discussing it.) P moves the third of a triad up or down a half step, as in Figure 48b. R raises the fifth of a major triad up a whole step or the root of minor triad down a whole step, as in Figure 48c. The D (dominant) operation is derivative because it decomposes into L + R; for example, a C major triad moves by L to an E minor triad, and the latter moves by R to a G major triad. It is easy to transform any triad into any other triad using L, P, and R.

N-R theory has not developed a metric of chord distances, but the discussion in Cohn (1996) implies the simple solution of counting each application of a transformation as one distance unit. In terms of *TPS*, this is comparable to applications of the *j* operator in δ . There is no analogue to the *i* operator because of the absence of modulation in N-R theory. A plausible next step is to reinstate the *k* operator. The result is the rule in Figure 49.

A form of the basic space is needed to count instances of uncommon tones for k . Since N-R theory does not include scales or an internal hierarchy for the triad, there can be only two levels in its basic space, the triad and the chromatic collection out of which the triad is built. Figure 50 illustrates the structure of the space and the application of the N-R chord distance rule to the first progression in the chromatic version of Wagner's Grail theme, which is repeated in Figure 51. The transformations are $P + L + P$, that is, $Eb_{maj} \rightarrow Eb_{min} (= D\sharp_{min}) \rightarrow B_{maj} \rightarrow B_{min}$. All three pitch classes in B_{min} are new, so $k = 3$. The result is $\delta_{nr} = 6$. This process can be carried out sequentially for the rest of the harmonic progressions in the theme.

The nontriadic tones—F and A♭ in event 6, G♭ in Event 8—must also be accounted for. For this purpose

we directly import the surface tension rule (Figure 9). Attractions, however, must be recomputed using the two-tiered space in Figure 50. Let the anchoring strength of the chromatic level be 1 and that of the triad level be 2. The denominator in the N-R version of the harmonic attraction rule (Figure 14) equals δ_{nr} . The resultant α_{th} values do not call for multiplication by constant c .

Now all the factors of the tension model are transported into the N-R framework except for prolongational structure. However, as shown in Figures 41 and 44, a largely sequential interpretation of the chromatic Grail theme best fits the empirical data. Hence it is reasonable to test the N-R adaptation sequentially, at least for this passage. Figure 51 does this. The numbers by the arrows record values of δ_{nr} . In accordance with the treatment of α when a chord does not change, the double arrows between Events 2-4 and 4-7 indicate the continuation of δ_{nr} values until the next chord. This procedure permits a complete sequential analysis. As before, the surface dissonance numbers are added to the δ_{nr} numbers to give the tension (T_{seq}) values; and these numbers combine by multiple regression with the attraction numbers to yield an overall prediction of tension. The correlation with the data, shown in Figure 52, is very good: $R^2(2,6) = .86, p < .003, R^2_{adj} = .82$; $p(\text{attraction}) = .02, \beta = .46$; $p(\text{tension}) = .001, \beta = .85$. Note that again the β weight for attraction is less than that for tension.

This result is close to those of the sequential-hierarchical interpretations generated by the *TPS* model (Figures 40 and 42b). Why? First, in adapting the N-R framework to predict tonal tension, it has been necessary to import

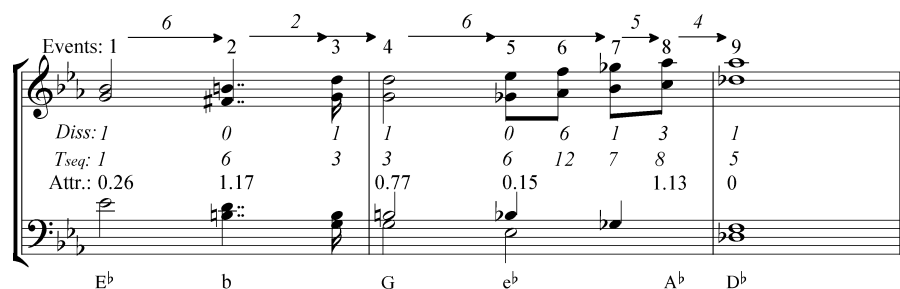


FIGURE 51. Neo-Riemannian tension analysis of the chromatic version of the Grail theme.

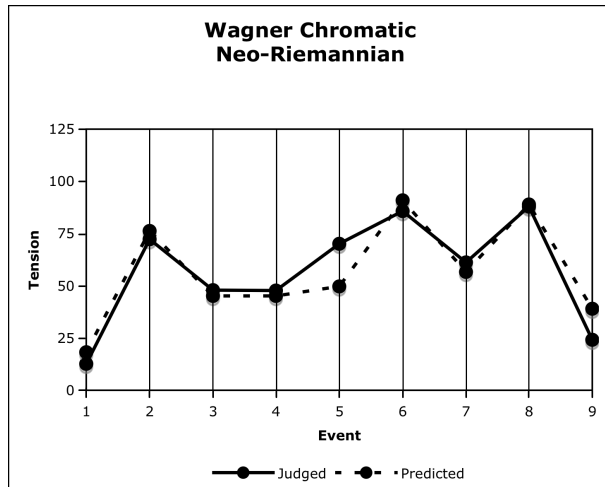


FIGURE 52. Tension graph for Figure 51.

much of the *TPS* model. Second, the hexatonic version of the *TPS* distance rule (δ_{hex}) employs equivalents of the L and P operations (R does not play a role in hexatonic space). P + L + P can be rephrased from “move the third of a major triad down a half step, then the fifth of the resultant minor triad up a half step, and then the third of the resultant major triad down a half step,” (thus $E_{b_{\text{maj}}} \rightarrow E_{b_{\text{min}}} \rightarrow B_{\text{maj}} \rightarrow B_{\text{min}}$) to “move the third of a major triad down one step on the hexatonic scale, then move each pitch class of the resultant minor triad down two steps on the hexatonic scale” ($E_{b_{\text{maj}}} \rightarrow E_{b_{\text{min}}} \rightarrow B_{\text{min}}$). Hook (2002) in effect takes these steps in a formalization of N-R theory.

The hexatonic distance rules in Chapter 6 of *TPS* perform the operations of the rephrased statement. The differences are that the *TPS* model includes a hexatonic modulation variable (i), assigns less weight to the chord transformation variable (j), and assigns more weight to

$$\text{TPS: } \delta_{\text{hex}}(E_{b/\text{hex}3} \rightarrow b/\text{hex}3) = 0 + 1 + 6 = 7$$

$$\text{N-R: } \delta_{\text{nr}}(E_{b \rightarrow b}) = \text{null} + 3 + 3 = 6$$

FIGURE 53. Comparison of *TPS* and N-R distance formulas for the progression $E_b \rightarrow b$.

the noncommon tone variable (k). The j variable has less weight because it reduces L and P to one expression. The k variable has more weight because *TPS* hexatonic space has more levels than does N-R hexatonic space, hence more instances of noncommon tones when a chord changes. The comparison can be seen for the progression $E_b \rightarrow b$ in Figure 53, in which δ_{hex} transposes the application of the *TPS* rule in Figure 36, and δ_{nr} modifies Figure 50 by putting a “null” placeholder for the absent i variable and combining L and P into a single j value. Despite the differences, the end results are almost the same.

Cohn ends his 1996 article by briefly considering how the LPR system might apply to diatonic music. In that spirit, Figure 54 carries out for the diatonic version of the Grail theme the same procedures that were applied to the chromatic version in Figure 51. Figure 55 plots the fit between predictions and data: $R^2(2,6) = .44$, $p = .18$, $R^2_{\text{adj}} = .26$; $p(\text{attraction}) = .55$, $\beta = .23$; $p(\text{tension}) = .19$, $\beta = .52$. Although the fit is poor, it is slightly better than the comparable sequential + attraction analysis, using the *TPS* methodology, in Figure 19. This may be because third-related progressions, which L and R yield, dominate the excerpt. A passage with many fifth-related progressions, which are statistically more common in diatonic music, would fare less well under a N-R approach. In any case, the weak correlation in Figure 54 reinforces from another perspective the necessity of a hierarchical analysis of diatonic music.

In conclusion, N-R transformations and distances approximate some of those of the *TPS* model. To accommodate predominantly diatonic music, however,

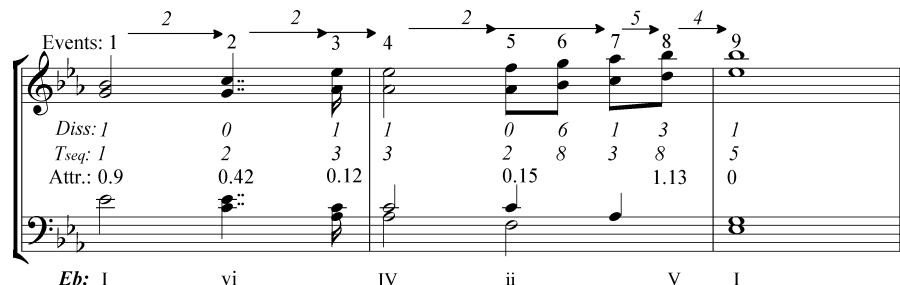


FIGURE 54. N-R tension analysis of the diatonic version of the Grail theme.

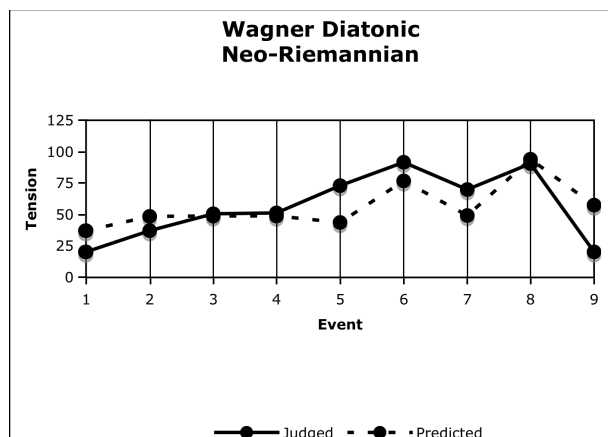


FIGURE 55. Tension graph for Figure 54.

the N-R approach must incorporate not only attractions but also a hierarchical tension analysis. Because it does not include any mechanism for modulation, the N-R approach is likely to be problematic for music that modulates within a space or across spaces. These issues deserve further empirical exploration. Balanced against empirical results are theoretical considerations such as the relative parsimony, expressivity, and generality of the *TPS* and N-R theories. We leave these complex matters here. Even this limited comparison has been useful to the extent that it suggests how a major trend in current music theory can be reworked to address an important psychological issue, the rise and fall of tonal tension.